OM nucleic - nucleic search, using sw model

```
                    GenCore version 5.1.9
          Copyright (c) 1993 - 2006  Bioacceleration Ltd.
```

Run on: June 14, 2006, 10:00:16 ; Search time 900.103 Seconds
                                            (without alignments)
                                            10248.037 Million cell updates/sec

Title: US-10-601-319-1
Perfect score: 1322
Sequence: 1 atgaaagcgatcttaatccc.........atcaccaccaccactaa 1323

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 348612423l residues

Total number of hits satisfying chosen parameters: 10489840

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Database :   N_Geneseq_8:*
                1: geneseqn1908:*
                2: geneseqn1908:*
                3: geneseqn2006s:*
                4: geneseqn2001as:*
                5: geneseqn2001bs:*
                6: geneseqn2002as:*
                7: geneseqn2002bs:*
                8: geneseqn2003as:*
                9: geneseqn2003bs:*
                10: geneseqn2003cs:*
                11: geneseqn2004as:*
                12: geneseqn2004aas:*
                13: geneseqn2004bs:*
                14: geneseqn2005s:*
                15: geneseqn2006s:*
```

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1322 | 100.0 | 1323 | 5 | AAC88885 | Aac88885 Escherich |
| 2 | 1322 | 100.0 | 1323 | 6 | AAD36473 | Aad36473 Escherich |
| 3 | 1322 | 100.0 | 1323 | 6 | AAD25460 | Aad25460 Escherich |
| 4 | 1322 | 100.0 | 1323 | 8 | ADA19445 | Ada19445 E. coli B |
| 5 | 1322 | 100.0 | 1323 | 12 | ADO50291 | Ado50291 Escherich |
| 6 | 1322 | 100.0 | 1323 | 15 | AEE75412 | Aee75412 E. coli B |
| 7 | 1315.6 | 99.5 | 1323 | 12 | ADO50301 | Ado50301 Kangaroo |
| 8 | 1315.6 | 99.5 | 1323 | 12 | ADO50303 | Ado50303 Escherich |
| 9 | 1287.8 | 97.4 | 1901 | 6 | AAD25463 | Aad25463 Escherich |
| 10 | 1287.8 | 97.4 | 1901 | 8 | ADA19449 | Ada19449 E. coli K |
| 11 | 1287.8 | 97.4 | 1901 | 10 | ADC87744 | Adc87744 DNA encod |
| 12 | 1287.8 | 97.4 | 1901 | 12 | ADO50297 | Ado50297 Escherich |
| 13 | 1287.8 | 97.4 | 1901 | 14 | AED50819 | Aed50819 Escherich |
| 14 | 1287.8 | 97.4 | 1901 | 15 | AEE75418 | Aee75418 E. coli B. |
| 15 | 1287 | 97.4 | 1299 | 3 | ABK12514 | Abk12514 DNA encod |
| 16 | 1287 | 97.4 | 1299 | 8 | ACA19297 | Aca19297 Prokaryot |
| 17 | 1287 | 97.4 | 1299 | 12 | ADL16137 | Adl16137 Escherich |
| 18 | 1284.6 | 97.2 | 1901 | 12 | ADO50295 | Ado50295 Escherich |
| 19 | 1284.6 | 97.2 | 1901 | 15 | AEE75416 | Aee75416 E. coli p |
| 20 | 1283.8 | 97.1 | 1901 | 12 | ADL16139 | Adl16139 Shigella |
| 21 | 1283 | 97.0 | 1901 | 12 | ADO50296 | Ado50296 Escherich |
| 22 | 1283 | 97.0 | 1283 | 15 | AEE75417 | Aee75417 E. coli p |
| 23 | 1282.2 | 97.0 | 3470 | 3 | AAC68298 | Aac68298 R15/APPA |
| 24 | 1282.2 | 97.0 | 4060 | 3 | AAC68298 | Aac68298 R15/APPA |
| 25 | 1282.2 | 97.0 | 6116 | 3 | AAC68297 | Aac68297 R15/APPA |
| 26 | 1282.2 | 97.0 | 6708 | 3 | AAC68295 | Aac68295 R15/APPA |
| 27 | 1282.2 | 97.0 | 17732 | 3 | AAC68300 | Aac68300 Lama2/APP |
| 28 | 1282.2 | 97.0 | 20623 | 3 | AAC68294 | Aac68294 Lama2/APP |
| 29 | 1281.4 | 96.9 | 1315 | 10 | ADL02197 | Adl02197 DNA encod |
| 30 | 1280.6 | 96.9 | 1296 | 12 | ADL16138 | Adl16138 Shigella |
| 31 | 1274.2 | 96.4 | 1299 | 15 | AEF15218 | Aef15218 Escherich |
| 32 | 1270.2 | 96.1 | 5421 | 3 | AAC68299 | Aac68299 SV40/APPA |
| 33 | 1267.8 | 95.9 | 1299 | 13 | ADW76362 | Adw76362 Phytase A |
| 34 | 1265 | 95.7 | 1308 | 10 | ADC87742 | Adc87742 DNA encod |
| 35 | 1265 | 95.7 | 1308 | 12 | ADO50299 | Ado50299 Escherich |
| 36 | 1265 | 95.7 | 1308 | 15 | AEE75420 | Aee75420 E. coli B |
| 37 | 1263.8 | 95.6 | 1489 | 3 | AAA28216 | Aaa28216 E. coli a |
| 38 | 1263.8 | 95.6 | 1489 | 8 | ACC57672 | Acc57672 E. coli a |
| 39 | 1263 | 95.5 | 1486 | 3 | AAD06831 | Aad06831 E. coli a |
| 40 | 1262.2 | 95.5 | 1486 | 8 | ADA19452 | Ada19452 E. coli K |
| 41 | 1257.4 | 95.1 | 1901 | 15 | AEE75422 | Aee75422 E. coli B |
| 42 | 1255 | 94.9 | 1296 | 14 | AED50817 | Aed50817 Escherich |
| 43 | 1251.8 | 94.7 | 1486 | 4 | AAD06832 | Aad06832 E. coli a |
| 44 | 1251.8 | 94.7 | 1486 | 8 | ACC57673 | Acc57673 Escherich |
| 45 | 1217 | 92.1 | 1304 | 14 | AED50824 | Aed50824 Escherich |

ALIGNMENTS

```
RESULT 1
AAC88885

ID    AAC88885 standard; DNA; 1323 BP.
XX
AC    AAC88885;
XX
DT    07-MAR-2001  (first entry)
XX
DE    Escherichia coli B phytase enzyme nucleotide sequence.
XX
KW    Escherichia coli B; phytase enzyme; anabolic; phytate digestion;
KW    nutrition; ds.
XX
OS    Escherichia coli.
XX
XX
XX
PN    WO200071728-A1.
XX
PD    30-NOV-2000.
XX
PF    25-MAY-2000; 2000WO-US014846.
XX
PR    25-MAY-1999;  99US-00318528.
XX
PA    (DIVE-) DIVERSA CORP.
XX
PI    Short JM,  Kretz KA;
XX
DR    WPI; 2001-112081/12.
DR    P-PSDB; AAB37892.
XX
PT    Improving the nutritional value of phytate-containing foodstuffs,
PT    phytase enzymes which catalyze the liberation of inorganic phosphate from
PT    the phytates.
XX
PS    Claim 2; Fig 1; 147pp; English.
XX
CC    The present sequence encodes a phytase enzyme from Escherichia coli B.
CC    The enzyme catalyses the liberation of inorganic phosphate from the
CC    phytate in phytate-containing foodstuffs and can thus be used to improve
CC    the nutritional value of phytate rich ingredients
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

```
SQ   Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;

Query Match             100.0%;  Score 1322;  DB 5;  Length 1323;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1323;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTTCGATTCGTTAACCCCCGAAATCTGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTTCGATTCGTTAACCCCCGAAATCTGCA 60

Qy         61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGTGGATTGTTCAGTGGTCGTCGTCATGGT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGTGGATTGTTCAGTGGTCGTCGTCATGGT 120

Qy        121 GTGGCGTGCTCCAAACCAAGGCCACGCAACTGATGCAGGAGTGCACCCCGACCGCATGGCCA 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTGGCGTGCTCCAAACCAAGGCCACGCAACTGATGCAGGAGTGCACCCCGACCGCATGGCCA 180

Qy        181 ACCTGGCCGGTTAAAACTGGGTTGGCTGACACCGGCGNGGTGGTGAGCGTAATCGCCTATCTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCTGGCCGGTTAAAACTGGGTTGGCTGACACCGGCGNGGTGGTGAGCGTAATCGCCTATCTC 240

Qy        241 GGACAATTACCAACGCCAGCGTCTGGGTAGCCGGACGGATTGCTGGCGAAAAAGGGCTGCCCG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGACAATTACCAACGCCAGCGTCTGGGTAGCCGGACGGATTGCTGGCGAAAAAGGGCTGCCCG 300

Qy        301 CAGTCTCGGTCAGGTCGCGGATTATTGCTGATGTCGACGAGCGGCGTACCCGTAAAACAGGCGGAA 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGTCTCGGTCAGGTCGCGGATTATTGCTGATGTCGACGAGCGGCGTACCCGTAAAACAGGCGGAA 360

Qy        361 GCCTTCGGCCGCGGGCGTGGCACCTGACTGTGCAATAACCGTACATACCCAGCGCAGATACG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTTCGGCCGCGGGCGTGGCACCTGACTGTGCAATAACCGTACATACCCAGCGCAGATACG 420

Qy        421 TCCAGTCCGGATCCGTTATTTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCAGTCCGGATCCGTTATTTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG 480

Qy        481 AACGTGACTGAGCGGATCCTCAGCAGGGCAGAGGGGTCAATTGCTGACTTTACCGGGCCAT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACGTGACTGAGCGGATCCTCAGCAGGGCAGAGGGGTCAATTGCTGACTTTACCGGGCCAT 540

Qy        541 CGGCAAACGGGGTTTGGCGGAACTGGGTGCTTAATTTTCCGCAATCAAAACTTGTGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGGCAAACGGGGTTTGGCGGAACTGGGTGCTTAATTTTCCGCAATCAAAACTTGTGC 600

Qy        601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC 660

Qy        661 AAGGTGAGCGCGACAAATGTCTCTCATTAACCGGTGGGTAAGCCTCGCATCAAATGCTGACG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGGTGAGCGCGACAAATGTCTCTCATTAACCGGTGGGTAAGCCTCGCATCAAATGCTGACG 720

Qy        721 GAGATATTTCTCCTGCAACAAGCACACAGGGAATGCCGGAGCGCGGGGTGGGGAGGATCACC 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGATATTTCTCCTGCAACAAGCACACAGGGAATGCCGGAGCGCGGGGTGGGGAGGATCACC 780

Qy        781 GATTCACCACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCCAATTTATTTGCTACACA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GATTCACCACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCCAATTTATTTGCTACACA 840

Qy        841 CGCACCGCCGAGGTTGCCCGCGGCCCACCCCGGTTATTCGATTGATTGATCATGGCAGCG 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CGCACCGCCGAGGTTGCCCGCGGCCCACCCCGGTTATTCGATTGATTGATCATGGCAGCG 900

Qy        901 TTGACCGCCCATCCACCGCAAAAACAGGCGTANGGGTGTGACATTACCCACTTCAGTACTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTGACCGCCCATCCACCGCAAAAACAGGCGTANGGGTGTGACATTACCCACTTCAGTACTG 960

Qy        961 TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGGCCACTGGAGCTCAACTGG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGGCCACTGGAGCTCAACTGG 1020

Qy       1021 ACGGCTTCCCGGGTGCAGCCGGATAACCGGCGCCAGGTGGTGGTGAACTGGTGGTTTGAACGCTCGG 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACGGCTTCCCGGGTGCAGCCGGATAACCGGCGCCAGGTGGTGGTGAACTGGTGGTTTGAACGCTCGG 1080

Qy       1081 CGTCGGGCTAAGCGATAACAGCCAGTGGAGAATTCAGGTTTCGTGGTTCTTCCAGACTTTACAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CGTCGGGCTAAGCGATAACAGCCAGTGGAGAATTCAGGTTTCGTGGTTCTTCCAGACTTTACAG 1140

Qy       1141 CAGATGCGTGATAAAAATACGCGCCGCCCGGAGAGGTGAAACTGACC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CAGATGCGTGATAAAAATACGCGCCGCCCGGAGAGGTGAAACTGACC 1200

Qy       1201 CTGGCAGGAGTGTGAAAGAGCGGAAATGCCGCAGGGGCATGTTTCCGTTGGCAGGGTTTTACGGCAA 1260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CTGGCAGGAGTGTGAAAGAGCGGAAATGCCGCAGGGGCATGTTTCCGTTGGCAGGGTTTTACGGCAA 1260

Qy       1261 ATCGTGAATGAAGCACCGGATACCGGGCGGCAGTTTGAGATCTCATCACCATCACCATCAC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ATCGTGAATGAAGCACCGGATACCGGGCGGCAGTTTGAGATCTCATCACCATCACCATCAC 1320

Qy       1321 TAA 1323
              |||
Db       1321 TAA 1323
```

RESULT 2
AAD36473
ID   AAD36473 standard; DNA; 1323 BP.
XX
AC   AAD36473;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Escherichia coli phytase DNA.
XX
KW   Dietary aid; biocompatible composition; therapeutic; digestive tract;
KW   foodstuff; digestion; phytase; enzyme; gene; ds.
XX
OS   Escherichia coli.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1323
FT   /tag= a
FT   /product= "E. coli phytase protein"
XX
PN   WO200189317-A2.
XX
PD   29-NOV-2001.
XX
PF   15-MAY-2001; 2001WO-US015764.
XX
PR   25-MAY-2000; 2000US-00580937.
XX
PA   (DIVE-) DIVERSA CORP.
XX
PI   Short JM, Kretz KA, O'donoghue E;
XX
DR   WPI; 2002-164149/21.
DR   P-PSDB; AAB22836.
XX
PT   New dietary aids comprising sustained release biocompatible compositions,
PT   comprise agent that assists in digestion, useful for delivering enzymes,
PT   therapeutics, medicine or agents to an organism.
XX
PS   Disclosure; Fig 1; 89pp; English.
XX
CC   The present invention relates to novel dietary aids comprising sustained
CC   release biocompatible composition which comprises an agent (enzymes such
CC   as phytase, amylase, esterase, protease) that assists in digestion. The
CC   biocompatible composition is effective upon oral consumption and release
CC   in the digestive tract of a subject. The dietary aids are useful for
CC   delivering enzymes, therapeutics, medicine and agents to an organism. The

```
CC   use of enzymes and other agents in digestive aids of livestock or
CC   domesticated animals not only improves the animal's health and life
CC   expectancy but also assists in increasing the health of livestock or in
CC   the production of foodstuffs from livestock. The present sequence is
CC   Escherichia coli phytase DNA
XX
SQ   Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;

Query Match             100.0%;  Score 1322;  DB 6;  Length 1323;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1323;   Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
Qy     1 ATGAAAAGCGATCTTAAATCCCATATTTATCTCTCTGATCTCCGTTAACCCCGCAAATCTGCA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAAAAGCCGATCTTAAATCCCATTTTATCTCTCTGATCTCCGTTAACCCCGCAAATCTGA 60

Qy    61 TTCGGCTCAGAGTGAGCCCGGAGCTGAAAGCTGGAGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TTCGGCTCAGAGTGAGCCCGGAGCTGAAAGCTGGAGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120

Qy   121 GTGCGTGCTCCAACCAAGGCCACCCAGCTGATGCAGGATGTCACCCCAGACGCATGGCCA 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTGCGTGCTCCAACCAAGGCCACCCAGCTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy   181 ACCTGGCCGGTAAAACTGGGTTGGTTGCTGACACCCCGCGGTTGTGAGCTAATCGGCCTATCTC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ACCTGGCCGGTAAAACTGGGTTGGTTGCTGACACCCCGCGGNGTGGTGTGAGCTAATCGGCCTATCTC 240

Qy   241 GGACATTACCAACGCCAGGCGTCTCGGTAGCCGACCGGATTGCTGGGCGAAAAAGGGCTGCCCG 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGACATTACCAACGCCAGCGTCTCGGTAGCCGACCGGATTGCTGGGCGAAAAAGGGCTGCCCG 300

Qy   301 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGACGTACCCGTTAAAAACAGGGGAA 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGACGTACCCGTTAAAAACAGGGGAA 360

Qy   361 GCCTTCGCGCCGCCGGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCCTTCGCGCCGCCGGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420

Qy   421 TCCAGTCCCGATCCGGTTATTAATCCCTAAAAAACTGGCGTTTGCCAACTGGATAACGCGG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCCAGTCCCGATCCGGTTATTAATCCCTAAAAAACTGGCGTTTGCCAACTGGATAACGCGG 480

Qy   481 AACGTGACTGACCGCGATCCTCAGCAGGGCGAGGGGTCAAATTCGTGACTTTACCGGGCAT 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AACGTGACTGACCGCGATCCTCAGCAGGGCGAGGGGTCAAATTCGTGACTTTACCGGGCAT 540

Qy   541 CGGGCAAACGGCGTTTCCGGAACTGGGACACGGGTGCTTAATTTCCGCAAATCAAACTTGTGC 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CGGGCAAACGGCGTTTCCGGAACTGGGACACGGGTGCTTAATTTCCGCAAATCAAACTTGTGC 600

Qy   601 CTTAAACGTGAGAAACAGGACAGAAGCTTGTTCATTAACGGACTCAGGCCATTACCATGGAACTC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTTAAACGTGAGAAACAGGACAGAAGCTTGTTCATTAACGGACTCAGGCCATTACCATGGAACTC 660

Qy   661 AAGGTGAGCGCCCAACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AAGGTGAGCGCCCAACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 720

Qy   721 GAGATATTTCTCTGCAACAACAGCACAGGGAAATGCCCGAGCCGGGGAAGGGATCACC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GAGATATTTCTCTGCAACAACAGCACAGGGAAATGCCCGAGCCGGGGAAGGGATCACC 780

Qy   781 GATTCACACCAGTGGAACACCTTGCTAAGGTTTGCATAACGGCAATTTTATTTGCTACACAA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GATTCACACCAGTGGAACACCTTGCTAAGGTTTGCATAACGGCAATTTTATTTGCTACACAA 840

Qy   841 CGCACGCCAGAGGTTGCCCGCAGCGCGCCACCCCGTTATTGGATTTGATTCATGGCACGCG 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CGCACGCCAGAGGTTGCCCGCAGCGCGCCACCCCGTTATTGGATTTGATTCATGGCACGCG 900

Qy   901 TTGACGCCCCATCCACCGCGAAAAACAGCGTATAGGGTGTGACATTACCCACTTCAGTACTG 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTGACGCCCCATCCACCGCGAAAAACAGCGTATGGGTGTGACATTACCCACTTCAGTACTG 960

Qy   961 TTTATTGCCGGACACGATACTAATCTGGCAAAATCTCGGCGGCCGCCACTGGAGCTCAACTGG 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TTTATTGCCGGACACGATACTAATCTGGCAAAATCTCGGCGGCGCCACTGGAGCTCAACTGG 1020

Qy  1021 ACGGCTTCCCGGTCAGCCGGATAACACCGCCCGCCAGGTGGTGGTGAACTGGTGTTTGAACGGCTGG 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 ACGGCTTCCCGGTCAGCCGGATAACACCGCCCGCCAGGTGGTGGTGAACTGGTGTTTGAACGGCTGG 1080

Qy  1081 CGTCCGGCTAAGCGATAACACCCAGTCGGATTCAGGTTTCGGTGGTTCTTCCAGACTTTACAG 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 CGTCCGGCTAAGCGATAACACCCAGTCGGATTCAGGTTTCGGTGGTTCTTCCAGACTTTACAG 1140

Qy  1141 CAGATGCGTGATAAAACGCGCTGTCATTAAAATAACGCCCCGCCCGGAGGGTGAAAACTGACC 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 CAGATGCGTGATAAAACGCGCTGTCATTAAAATAACGCCCCGCCCGGAGGGTGAAAACTGACC 1200

Qy  1201 CTGGCAGGATGTGAAGAGCCGAAAATCGGCAGGGCCATGGTTTCGGTTTTACGGCAA 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 CTGGCAGGATGTGAAGAGCCGAAAATCGGCAGGGCCATGGTTTCGGTTTTACGGCAA 1260

Qy  1261 ATCGTGAATGAACGACAGCGCATACCGGCGTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 ATCGTGAATGAACGACAGCGCATACCGGCGTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320

Qy  1321 TAA 1323
         |||
Db  1321 TAA 1323
```

RESULT 3
AAD25460

```
ID   AAD25460   standard; DNA; 1323 BP.
XX
AC   AAD25460;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Escherichia coli B phytase DNA.
XX
KW   Bacterial phytase; K12 appA phytase; protease stability; anabolic;
KW   gastrointestinal; nutritional value; feed treatment process; therapy;
KW   thermal tolerance; growth performance; alcoholic drink; biopulping;
KW   non-alcoholic drink; biobleaching; B phytase; ds.
XX
OS   Escherichia coli.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1323
FT                   /tag= a
FT                   /product= "E. coli B phytase protein"
XX
PN   WO20019033-A2.
XX
PD   29-NOV-2001.
XX
PF   24-MAY-2001; 2001WO-US017118.
XX
PR   25-MAY-2000; 2000US-00580515.
XX
PA   (DIVE-) DIVERSA CORP.
XX
PI   Short JM,  Kretz KA,  Gray KA,  Barton NR,  Garrett JB,  O'donoghue E;
XX
DR   WPI; 2002-083108/11.
DR   P-PSDB; AAB15806.
XX
PT   New bacterial phytase for e.g. improving the nutritional value of phytate
PT   -containing foodstuffs and subsequently improving the growth performance
PT   of an organism that consumes it, or in treating animal digestive systems.
XX
```

```
PS
XX
CC   The patent discloses recombinant bacterial phytase from Escherichia coli
CC   K12 appA phytase. The enzyme has phytase activity and improved thermal
CC   tolerance when compared with wild-type phytase. It has improved protease
CC   stability at low pH. The recombinant phytase is useful for improving the
CC   nutritional value of phytate-containing foodstuffs and subsequently
CC   improving the growth performance of an organism that consumes it, in
CC   treating animal digestive systems, in feed treatment processes and for in
CC   vitro purposes related to research, discovery and development. They are
CC   also used for generating recombinant digestive system life forms, for
CC   producing or manufacturing alcoholic and non-alcoholic drinks based on
CC   the use of moulds, grains and/or plants, in biopulping and bio-bleaching
CC   where a reduction in the use of environmentally harmful chemicals that
CC   are traditionally used in the pulp and paper industry is desired and in
CC   the reduction or possible elimination of the need for mineral
CC   supplements, enzymes or therapeutic drugs for animals present in the daily feed
CC   thus increasing the amount calories and nutrients present in the feed.
CC   The present sequence is a DNA encoding E. coli B phytase protein
XX
SQ   Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;

Query Match             100.0%;  Score 1322;  DB 6;  Length 1323;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy       1 ATGAAAACGGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCGCGCAATCTGCA 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGAAAACGGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCGCGCAATCTGCA 60

Qy      61 TTCGCTTCAGAGTGAGCCGGGAGCTGAAGCTGGAAAAGTGTGGTGATATTGTCAGTCATGGT 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TTCGCTTCAGAGTGAGCCGGGAGCTGAAGCTGGAAAAGTGTGGTGATATTGTCAGTCATGGT 120

Qy     121 GTGGCGTGCTCCAACCAAGGCCACGCACGCAACTGATGTCGCAGGANTGTCACCCCGACGCCA 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GTGGCGTGCTCCAACCAAGGCCACGCACGCAACTGATGTCGCAGGANTGTCACCCCGACGCCA 180

Qy     181 ACCTGGCGCGGTAAAACTGGGTTGGCTGACACCCGCGNGGTGGTGAGCTAATCGGCCTATCTC 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ACCTGGCGCGGTAAAACTGGGTTGGCTGACACCCGCGNGGTGGTGAGCTAATCGGCCTATCTC 240

Qy     241 GGACANTTACCAAACGCCAGCGTCGTCTGGTAGCCCACGCGAGTTGCTTCGGCGAAAAAGGGC 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GGACANTTACCAAACGCCAGCGTCGTCTGGTAGCCCACGCGAGTTGCTTCGGCGAAAAAGGGC 300

Qy     301 CAGTCTGGTTCAGGTCGGGCGGATTATTNGCTCAATAACCGTACANACTTGTGC 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CAGTCTGGTTCAGGTCGGGCGGATTATTNGCTCAATAACCGTACANACTTGTGC 360

Qy     361 GCCTTCCCGCGCCGGGCGTGCACCTGACTGTGCAATAACCGTACACCCAGGCAGATACG 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GCCTTCCCGCGCCGGGCGTGCACCTGACTGTGCAATAACCGTACACCCAGGCAGATACG 420

Qy     421 TCCAGTCCCGATCCGTTATTTATTTAATCCTCTAAAAAACTGGGCGTTTGCCAACTGGATAACGCG 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TCCAGTCCCGATCCGTTATTTATTTAATCCTCTAAAAAACTGGGCGTTTGCCAACTGGATAACGCG 480

Qy     481 AACGTGACTGACGCGGATCCTCAGCAGGGCAGAGGGTCAATTGCTGACTTTACCGGGCAT 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AACGTGACTGACGCGGATCCTCAGCAGGGCAGAGGGTCAATTGCTGACTTTACCGGGCAT 540

Qy     541 CGGCAAACGGCGTTTCCGGAACTGGACGGGGTGCTTAATTTTCCGGCAATCAAAACTTGTGC 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 CGGCAAACGGCGTTTCCGGAACTGGACGGGGTGCTTAATTTTCCGGCAATCAAAACTTGTGC 600

Qy     601 CTTAAAACGTGAGAAAACAGGACGCAGGCACTATACCATCGGAACTC 660
           |||||||||||||||||||||||||||||||||||||||||||||||
Db     601 CTTAAAACGTGAGAAAACAGGACGCAGGCACTATACCATCGGAACTC 660

Qy     661 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTAAGCCTCGGCATCAATGCTGACG 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTAAGCCTCGGCATCAATGCTGACG 720
```

```
Qy     721 GAGATATTTCTCCTGCCAACAAGCACAGGGAATGCCGGAGCCGGGGGTGGGGAGGATCACC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GAGATATTTCTCCTGCCAACAAGCACAGGGAATGCCGGAGCCGGGGGTGGGGAGGATCACC 780

Qy     781 GATTCACACCAGTGGACAACCTTGCTAAGTTTGCATAAGCGCGCAATTTTATTTGCTACAA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GATTCACACCAGTGGACAACCTTGCTAAGTTTGCATAAGCGCGCAATTTTATTTGCTACAA 840

Qy     841 CGCACGCCGAGAGTTGCCCGCGCCAGCCGCCACCCGTTATTGATTTGATTTGATCANTGGCAGCG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CGCACGCCGAGAGTTGCCCGCGCCAGCCGCCACCCGTTATTGATTTGATTTGATCANTGGCAGCG 900

Qy     901 TTGACGCCCCATCCACCCGCAAAAACAGGCGTANGGGTGGTGGACATTACCCACTTCAGTACTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 TTGACGCCCCATCCACCCGCAAAAACAGGCGTANGGGTGGTGGACATTACCCACTTCAGTACTG 960

Qy     961 TTTATTGCCGGACACGATACTAATCTGGCAAAATCTCGGCGCGCCACTGGAGCTCAACTGG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 TTTATTGCCGGACACGATACTAATCTGGCAAAATCTCGGCGCGCCACTGGAGCTCAACTGG 1020

Qy    1021 ACGGCTTCCCGGTCAGCCGGATAACACCCGCCCAGGGTTGTGAACTGGGCGTCG 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 ACGGCTTCCCGGTCAGCCGGATAACACCCGCCCAGGGTTGTGAACTGGGCGTCG 1080

Qy    1081 CGTCGGCTNAAGCCGATAACGACCCAGTGGATTCAGGTTNGGCTGGTCGTTCTTCCAGACTTTACAG 1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CGTCGGCTNAAGCCGATAACGACCCAGTGGATTCAGGTTNGGCTGGTCGTTCTTCCAGACTTTACAG 1140

Qy    1141 CAGATGCCTGATAAAAACGCGCTGTCATTAAAATACGCGCCCCGGAGAGGTGAAACTGACC 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 CAGATGCCTGATAAAAACGCGCTGTCATTAAAATACGCGCCCCGGAGAGGTGAAACTGACC 1200

Qy    1201 CTGGCAGGATGTGAAGAGCGAAATGGCGCAGGGCATGTTGTTCGGTTGCAGGTTTTACGGCAA 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 CTGGCAGGATGTGAAGAGCGAAATGGCGCAGGGCATGTTGTTCGGTTGCAGGTTTTACGGCAA 1260

Qy    1261 ATCGTGAATGAAGCAACGGCATACGGGGCCTGGCAGTTTGAGATCTCATCACCATCACCATCAC 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 ATCGTGAATGAAGCAACGGCATACGGGGCCTGGCAGTTTGAGATCTCATCACCATCACCATCAC 1320

Qy    1321 TAA 1323
           |||
Db    1321 TAA 1323
```

```
RESULT 4
ADA19445
ID   ADA19445 standard; DNA; 1323 BP.
XX
AC   ADA19445;
XX
DT   20-NOV-2003 (first entry)
XX
DE   E. coli B DNA encoding phytase.
XX
KW   Phytase; ds; gene; phytate; appA gene; animal feed; inorganic phosphate;
KW   digestion enhancement; transgenic; thermal tolerance; protease stability.
XX
OS   Escherichia coli B.
XX
FH   Key              Location/Qualifiers
FT   CDS              1..1323
FT                    /*tag= a
FT                    /product= "Phytase"
XX
PN   US2002136754-A1.
XX
PD   26-SEP-2002.
XX
PF   24-MAY-2001; 2001US-00866379.
XX
PR   13-AUG-1997; 97US-00910798.
```

```
PR   01-MAR-1999;
PR   (SHOR/)   SHORT J M.
PA   (KRET/)   KRETZ K A.
PR   13-APR-1999;   99US-00259214.
PR   25-MAY-1999;   99US-00291931.
PR   25-MAY-2000;   2000US-00580515.
XX
PA   (SHOR/)   SHORT J M.
PA   (KRET/)   KRETZ K A.
PA   (GRAY/)   GRAY K A.
PA   (BART/)   BARTON N R.
PA   (GARR/)   GARRETT J B.
PA   (DONO//)  O' DONOGHUE E.
PA   (MATH//)  MATHUR E J.
XX
PI   Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB;
PI   O' Donoghue E, Mathur EJ;
XX
PS   Claim 2; Fig 1; 62pp; English.
XX
PT   The invention relates to an isolated Escherichia coli polynucleotide
PT   encoding a phytase enzyme appearing as ADA19450 and having amino acids
PT   modification from W68E, Q84W, A95P, K97C, S168E, R180Y, N226C or Y277D.
PT   Also included the E. coli appA gene ADA19449 (or an oligonucleotide
PT   derived from it) or its mutant sequence ADA19452, expression vectors,
PT   host cells, a method of improving nutritional value of a phytate-
PT   containing foodstuff by contacting the phytate-containing foodstuff with
PT   a pure phytase enzyme mutated phytase (where the phytase enzyme catalyzes
PT   the liberation of inorganic phosphate from the phytate in the phytate-
PT   containing foodstuff), a method to produce an animal feed containing a
PT   microbial phytase (comprising culturing the plant cell, plant part or
PT   plant under conditions where the nucleotide sequence is expressed and
PT   converting the plant cells, plants or plant into a composition for animal
PT   feed), a feed composition for animals (comprising the plant seeds, plant
PT   cells, plant parts or plants in admixture with a phytate-containing
PT   foodstuff), a method to treat a human or an animal able to benefit from
PT   digestive enhancement by the activity of an exogenous phytase enzymes
PT   comprising administering to the human or animal the plant seed, plant
PT   cells, plant parts or plants of a transgenic plant which is modified to
PT   contain an expression system which expresses a nucleotide sequence
PT   encoding a phytase enzyme, a transgenic non-human organism whose genome
PT   comprising a heterologous nucleic acid sequence encoding a polypeptide
PT   having phytase activity. The phytase enzyme is useful for improving the
PT   nutritional value of phytate-containing foodstuff, in the production of
PT   animal feed and for enhancing digestion in humans and animals. The
PT   invented method improves thermal tolerance and protease stability. It
PT   also improves the feeding value of phytate rich ingredients. The present
PT   invention encodes E. coli B wild-type phytase.
XX
SQ   Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;
```

The alignment block and match statistics (rotated text) include:

```
Query Match          100.0%;  Score 1322;  DB 8;  Length 1323;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 1323;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```
Db   1261 ATGCGTGAAATGAAGCACGCATACGGCGGCCTGCAGTTTGAGATCTTCATCACCATCACCATCAC 1320
Qy   1321 TAA 1323
          |||
Db   1321 TAA 1323
```

```
RESULT 5

ID   AD050291 standard; DNA; 1323 BP.
XX
AC   AD050291;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Escherichia coli B phytase DNA.
XX
KW   Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW   phytate; animal feed; fish feed; dough; baking; gene; ds.
XX
OS   Escherichia coli; B.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1323
FT                   /*tag= a
FT                   /product= "Phytase enzyme"
FT   unsure          214..216
FT                   /*tag= b
FT                   /note= "Encodes Arg"
XX
PN   US200409196 8-A1.
XX
PD   13-MAY-2004.
XX
PF   20-JUN-2003;  2003US-00601319.
XX
PR   13-AUG-1997;  97US-00910798.
PR   01-MAR-1999;  99US-0259214.
PR   13-APR-1999;  99US-00291931.
PR   25-MAY-1999;  99US-00318528.
PR   25-MAY-2000;  2000US-00580515.
PR   24-MAY-2001;  2001US-0086379.
XX
PA   (SHOR/)  SHORT J M.
PA   (KRETZ/)  KRETZ K.
PA   (GRAY/)  GRAY K A.
PA   (BART/)  BARTON N R.
PA   (GARR/)  GARRETT J B.
PA   (ODON/)  O'DONOGHUE E.
PA   (MATH/)  MATHER E J.
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI   Mather EJ;
XX
DR   WPI; 2004-374952/35.
DR   P-PSDB; AD050292.
XX
PT   Producing phytase, involves providing nucleic acid derived from bacteria
PT   encoding polypeptide having phytase activity, and expressing nucleic acid
PT   in yeast.
XX
PS   Claim 3; SEQ ID NO 1; 74pp; English.
XX
CC   The invention relates to a method for producing phytase that involves
CC   providing a nucleic acid encoding phytase derived from a bacteria, and
CC   expressing the nucleic acid in a yeast under conditions that allow
CC   expression of the enzyme in the yeast. The invention also relates to
CC   modified phytase enzyme which has improved thermal tolerance and protease
CC   stability at low pH. The phytase enzyme can be used in foodstuffs to
CC   improve the feeding value of phytate rich ingredients, and in diet of
CC   numerous animals including mammals, fowls and fishes, commercially
CC   significant mammals such as pigs, goats, laboratory rodents, commercially
CC   significant avian species such as chicken, ducks, doves, parrot, etc.,
CC   commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC   etc., in dough making and baking, in dietary aids for animals. The method
CC   provides easy manufacture of the active ingredient loaded biocompatible
CC   composition, higher yields and loading efficiency. The phytase
CC   incorporated in to the dietary aid is safe for animals. The present
CC   sequence is Escherichia coli B phytase DNA.
XX
SQ   Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;

Query Match             100.0%;  Score 1322;  DB 12;  Length 1323;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1323;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 ATGAAAGGGATCTTAATCCCATTTTTTATCTCTTCTGATTCCGTTAACCCCGCAAATCTGCA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAGGGATCTTAATCCCATTTTTTATCTCTTCTGATTCCGTTAACCCCGCAAATCTGCA 60

Qy     61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120

Qy    121 GTGGGTGGTCCAACCAAGGGCCACGCCAACTGATGCAGGAGTGTCACCCACCCCAGACCCCA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGGGTGGTCCAACCAAGGGCCACGCCAACTGATGCAGGAGTGTCACCCACCCCAGACCCCA 180

Qy    181 ACCTGGCCGGGTAAAACTGGGGTTGGCTGAACTGACACCCGCNGGGGTGGTGAGCTAATCTC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACCTGGCCGGGTAAAACTGGGGTTGGCTGAACTGACACCCGCNGGGGTGGTGAGCTAATCTC 240

Qy    241 GGACATTACCAACGCAGCGTCTGGTAGCCGACGGATTGCTGCTGGCGGAAAAAGGGCTGCCCG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGACATTACCAACGCAGCGTCTGGTAGCCGACGGATTGCTGCTGGCGGAAAAAGGGCTGCCCG 300

Qy    301 CAGTCTCGGTCAGGTCCGGATTAATTGCTGATGNGTCGACCGGCGTACCCCGTAAAACAGGGGAA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGTCTCGGTCAGGTCCGGATTAATTGCTGATGNGTCGACCGGCGTACCCCGTAAAACAGGGGAA 360

Qy    361 GCCTTCGGCGCCGGGGCTGGCACCTGACTGTGCAATAACGGTACATACCCAGGCAGATACG 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCTTCGGCGCCGGGGCTGGCACCTGACTGTGCAATAACGGTACATACCCAGGCAGATACG 420

Qy    421 TCCAGTCCCGATCCGTTTATTTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCCAGTCCCGATCCGTTTATTTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG 480

Qy    481 AACGTGACTGACGGCGATCCTCAGCACGGGCACGAGAGGGTCAATTGCTGACTTTACCGGGCAT 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AACGTGACTGACGGCGATCCTCAGCACGGGCACGAGAGGGTCAATTGCTGACTTTACCGGGCAT 540

Qy    541 CGGCAAACGGCGTTTCGGGAACGGGGTGCTTAATTTTCCGGCAATCAAACTTGTGC 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CGGCAAACGGCGTTTCGGGAACGGGGTGCTTAATTTTCCGGCAATCAAACTTGTGC 600

Qy    601 CTTAAACGTGAGAAAACAGGACGAAAAGCTGTTCATTAACGCAGCCGGAGCGGAACTC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTTAAACGTGAGAAAACAGGACGAAAAGCTGTTCATTAACGCAGCCGGAGCGGAACTC 660

Qy    661 AAGGTGAGCGCGCCGAACAAATGTCTCATTAACCGGTGTGCGGTAAGCCTTCGGCATCAATGCTCACG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AAGGTGAGCGCGCCGAACAAATGTCTCATTAACCGGTGTGCGGTAAGCCTTCGGCATCAATGCTCACG 720

Qy    721 GAGATATTTCTCCTGCAACAAGCAACAGGGAATGCCGGAGCCGGGGGTGGGGAAGGGGATCACC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GAGATATTTCTCCTGCAACAAGCAACAGGGAATGCCGGAGCCGGGGGTGGGGAAGGGGATCACC 780

Qy    781 GATTCACACCACTGGGAACACACCTTGCTAAGTTTGCATAAACGCGCAATTTTATTTTGCTACAA 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GATTCACACCACTGGGAACACACCTTGCTAAGTTTGCATAAACGCGCAATTTTATTTTGCTACAA 840

Qy    841 CGCACGCCAGAGCGTTGCCCGCACGACCCACCCCGTTATTCGGATTTGATCATGGCAGCG 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CGCACGCCAGAGCGTTGCCCGCACGACCCACCCCGTTATTCGGATTTGATCATGGCAGCG 900
```

```
Qy    901  TTGACGCGCCCATCCACCGCCGCAAAAACAGGGGCGTAGTCGGTGTGTGACATTACCCACTTCAGTACTG   960
           |||||| |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  TTGACGCCCCATCCACCGCCGCAAAAACAGGGGCGTAGTCGGTGTGTGACATTACCCACTTCAGTACTG   960

Qy    961  TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGGCGGCGCACTGGAGCTGAACTGG          1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGGCGGCGCACTGGAGCTGAACTGG          1020

Qy   1021  ACGCTTCCCGGTCAGCCCGGATAACACGCGCGCCAGGTGTGGAACTGGTGTTTGAACGGCTGG        1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  ACGCTTCCCGGTCAGCCCGGATAACACGCGCGCCAGGTGTGGAACTGGTGTTTGAACGGCTGG        1080

Qy   1081  CGTCGGCTAAGCGGATAACAGCCCAGTGGATTCAGGTTTCGCTGGTGCTTCCAGACTTTACAG       1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  CGTCGGCTAAGCGGATAACAGCCCAGTGGATTCAGGTTTCGCTGGTGCTTCCAGACTTTACAG       1140

Qy   1141  CAGATGCGTGATAAAAACGCGCGCTGTCATTAAATACGCGCCCGGAGAGGTGAAACTGACC         1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  CAGATGCGTGATAAAAACGCGCGCTGTCATTAAATACGCGCCCGGAGAGGTGAAACTGACC         1200

Qy   1201  CTGGCAGGATGTGAAGACGGGAAATGCGGCAGGCATGTGTTCGTTGGTGCAGGTTTTACGGCAA    1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  CTGGCAGGATGTGAAGACGGGAAATGCGGCAGGCATGTGTTCGTTGGTGCAGGTTTTACGGCAA    1260

Qy   1261  ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTGAGATCTCATCACCATCACCACCATCAC      1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261  ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTGAGATCTCATCACCATCACCACCATCAC      1320

Qy   1321  TAA 1323
           |||
Db   1321  TAA 1323
```

RESULT 6
AEE75412

```
ID   AEE75412 standard; DNA; 1323 BP.
XX
AC   AEE75412;
XX
DT   23-FEB-2006  (first entry)
XX
DE   E. coli B phytase coding sequence SEQ ID NO: 1.
XX
KW   ds; coding sequence; thermostable; phytase; protein stabilization;
KW   pharmaceutical; fertilizer; cramp; cell culture; osteoporosis;
KW   osteopathic; food.
XX
OS   Escherichia coli B.
XX
FH   Key             Location/Qualifiers
FH
FT   CDS             1..1323
FT                   /tag= a
FT                   /product= "Phytase"
FT   unsure          214..216
FT                   /tag= b
FT                   /note= "ecodes Arg"
XX
PN   US2005281792-A1.
XX
PD   22-DEC-2005.
XX
PF   01-SEP-2004; 2004US-00933115.
XX
PR   13-AUG-1997;  97US-00910798.
PR   01-MAR-1999;  99US-0259214.
PR   13-APR-1999;  99US-00291931.
PR   25-MAY-1999;  99US-00318528.
PR   25-MAY-2000;  2000US-00580515.
PR   24-MAY-2001;  2001US-00866379.
XX
PA   (SHOR/) SHORT J M.
PA   (KRET/) KRETZ K A.
```

Claim 1; SEQ ID NO 1; 82pp; English.

Formulation useful as dietary supplement for treating, preventing or
reversing osteoporosis and bone loss, and preventing muscle cramps,
comprises one or more polypeptides having phytase activity.

The present sequence is that of a polynucleotide encoding the Escherichia
coli B phytase. The present invention relates to a novel formulation,
comprising at least one polypeptide having phytase activity, where the
polypeptide is a modified sequence (AEE75421) derived from the
Escherichia coli K-12 appA phytase (AEE75419). The modification of the
enzyme, comprising a series of mutations, improves the thermal tolerance
and protease stability of the protein. The specification also claims a
pharmaceutical composition; a kit; an immobilized phytase; a fertilizer
or soil additive; a liquid supplement for preventing muscle cramps; a
hydrating agent; a tissue culture or cell culture media; and a plant food
additive, all comprising the novel phytase, and a method of reducing
pollution and increasing nutrient availability in an environment or
environmental sample by degrading environmental phytic acid. The novel
enzyme is a dietary supplement useful for treating, preventing or
reversing osteoporosis or bone loss, and preventing muscle cramps. The
liquid supplement is useful for reducing pollution and increasing muscle cramps. The method is
useful for reducing pollution and increasing nutrient availability in an
environment or environmental sample by degrading environmental phytic
acid, where the environment or environmental sample comprises a soil or a
body of water. The immobilized phytase is useful in foodstuffs for
improving the feeding value of phytate rich ingredients.

```
SQ   Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;

Query Match          100.0%;  Score 1322;  DB 15;  Length 1323;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 1323;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1  ATGAAAGCGATCTTAATCCCATTTTATCTCTTCTGATTCGGTAAACCCGCGAATCTGCA          60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  ATGAAAGCGATCTTAATCCCATTTTATCTCTTCTGATTCGGTAAACCCGCGAATCTGCA          60

Qy     61  TTCGCTCAGAGTGAGCCCGGACGCTGAAAGTGTGGAAAGTGTGTGATTGTCAGTCGTCATGGT       120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  TTCGCTCAGAGTGAGCCCGGACGCTGAAAGTGTGGAAAGTGTGTGATTGTCAGTCGTCATGGT       120

Qy    121  GTGGGTGCTTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA         180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  GTGGGTGCTTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA         180

Qy    181  ACCTGGCCGGTAAAAACTGGGTTGGCTGACACCGCGCGGNGGTGGAGCTAATCGCCTATCTC      240
           ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    181  ACCTGGCCGGTAAAAACTGGGTTGGCTGACACCGCGCGGNGGTGGAGCTAATCGCCTATCTC      240

Qy    241  GGACATTACCAAGCGCCAGCGGTTCTGGTAGCCGACGGAGTTGCTGGCGAAAAAGGCTGCCCG     300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  GGACATTACCAAGCGCCAGCGGTTCTGGTAGCCGACGGAGTTGCTGGCGAAAAAGGCTGCCCG     300

Qy    301  CAGTCTGGTCAGGTTCGCGCGATTATTGCTGATGTCACGACGTACCCGTAAAACAGGCGAA       360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  CAGTCTGGTCAGGTTCGCGCGATTATTGCTGATGTCACGACGTACCCGTAAAACAGGCGAA       360

Qy    361  GCCTTCGCGCGCGGCGGCGGTCGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG  420
```

```
Db    361 GCCTTCGCCCGCCCGGCTGGGCACCTGACTGTGTGCAATAACCGTACCCAGGCCAGATACG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    421 TCCAGTCCCGACCNCGGTTATTTATTTAATCTTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCCAGTCCCGACCNCGGTTATTTATTTAATCTTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG 480
Qy    481 AACGTGACTGACGGCGGATCCTCAGCACAGGGGGTCAAATTGCTGACTTTACCGGGGCAT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AACGTGACTGACGGCGGATCCTCAGCACAGGGGGTCAAATTGCTGACTTTACCGGGGCAT 540
Qy    541 CGGCAAACGGGCGTTTCGGCGGAACTGGGCGGGTGCTTAATTTTCCGCAATCAAACTTCGTGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CGGCAAACGGGCGTTTCGGCGGAACTGGGCGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600
Qy    601 CTTAAACGTGAGAAAACAGGACAGAAAGTCTGTTCATTAAACGCACAGGCATTACCATCGGAACTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTTAAACGTGAGAAAACAGGACAGAAAGTCTGTTCATTAAACGCACAGGCATTACCATCGGAACTC 660
Qy    661 AAGGTGAGCGCCGACAAATGTCTTCATTAACCGGTGCGGGTTAAGCCTCGGCATTACCGGGAACG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AAGGTGAGCGCCGACAAATGTCTTCATTAACCGGTGCGGGTTAAGCCTCGGCATTACCTGACG 720
Qy    721 GAGATATTTCTCCTGCAACAAGCACAGGGAAATCCGGAGACCCGGGTGGGGAAGGATCACC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GAGATATTTCTCCTGCAACAAGCACAGGGAAATCCGGAGACCCGGGTGGGGAAGGATCACC 780
Qy    781 GATTCACCAGTGGAACACCTTGCTTAGTGTTTGCATAACGCGCAAATTTATTTATTTGCTACAA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GATTCACCAGTGGAACACCTTGCTTAGTGTTTGCATAACGCGCAAATTTATTTATTTGCTACAA 840
Qy    841 CGCACGGCCAGAGAGGTTGCCCGCAGCCGCCCACCCCGGTTATTGGATTTGGATTCATGGCAGCG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CGCACGGCCAGAGAGGTTGCCCGCAGCCGCCCACCCCGGTTATTGGATTTGGATTCATGGCAGCG 900
Qy    901 TTGACGGCCCATCCACCCGCAAAAAACAGGCGTATGGGTGTGACATTACCCACTTCAGTACTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTGACGGCCCATCCACCCGCAAAAAACAGGCGTATGGGTGTGACATTACCCACTTCAGTACTG 960
Qy    961 TTTATTGCCGGACACAGCAGTGGATTCAGGTTCGCTGCTGTCTTCCAGACTTTACAG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTTATTGCCGGACACAGCAGTGGATTCAGGTTCGCTGCTGTCTTCCAGACTTTACAG 1020
Qy   1021 AGCGCTTCCCGGTCAGCCGGATAACACGCCGCCCGGAGGAGGTTGGATGGTTGAACGGCTGG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGCTTCCCGGTCAGCCGGATAACACGCCGCCCGGAGGAGGTTGGATGGTTGAACGGCTGG 1080
Qy   1081 CGTCGGCTAAGCCATAACAGCCAGTGGAATTCAGGTTCGCTGCTGTCTTCCAGACTTTACAG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ATCGTGAATGAAGCACCGCATACCGGCGTGCAGTTTGAGATCTCATCACCATCAC 1140
Qy   1141 CAGATGCCTGATGATAAAACCGGCTGTCATTAAAATACGGAGGTGAAACTGACC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAGATGCCTGATGATAAAACCGGCTGTCATTAAAATACGGAGGTGAAACTGACC 1200
Qy   1201 CTGGCAGGAGTGTGAAGAGCGAAAATGCCAAGGGCAATGTGTTCGTTGGGCAGGTTTTACGCAA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGCAGGAGTGTGAAGAGCGAAAATGCCAAGGGCAATGTGTTCGTTGGGCAGGTTTTACGCAA 1260
Qy   1261 ATCGTGAATGAAGCACAGCGATACCGGCGGTGCAGTTTGAGATCTCATCACCATCAC 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAGCACAGCGATACCGGCGGTGCAGTTTGAGATCTCATCACCATCAC 1320
Qy   1321 TAA 1323
          |||
Db   1321 TAA 1323
```

RESULT 7
ID   AD050301   standard; DNA; 1323 BP.
XX
AC   AD050301;
XX
DT   29-JUL-2004 (first entry)
XX
DE   Kangaroo rat Escherichia coli phytase 872PH1 mutant DNA.
XX
KW   Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW   phytate; animal feed; fish feed; dough; baking; mutant; gene; ds.
XX
OS   Escherichia coli.
XX
FH   Key   Location/Qualifiers
FH
FT   CDS   1..1323
FT   /*tag= a
FT   /product= "Phytase mutant enzyme"
FT   unsure   214..216
FT   /*tag= b
FT   /note= "Encodes Arg"
FT   mutation   replace(527,A)
FT   /*tag= c
FT   mutation   replace(893,T)
FT   /*tag= d
FT   mutation   replace(895,G)
FT   /*tag= e
FT   mutation   replace(934,G)
FT   /*tag= f
XX

CC   The invention relates to a method for producing phytase that involves
CC   providing a nucleic acid encoding phytase derived from a bacteria, and
CC   expressing the nucleic acid in a yeast under conditions that allow
CC   expression of the enzyme in the yeast. The invention also relates to
CC   modified phytase enzyme which has improved thermal tolerance and protease
CC   stability at low pH. The phytase enzyme can be used in foodstuffs to
CC   improve the feeding value of phytate rich ingredients, and in diet of
CC   numerous animals including mammals, fowls and fishes, commercially
CC   significant avian species such as chicken, ducks, doves, parrot, etc.,
CC   significant mammals such as pigs, goats, laboratory rodents, commercially
CC   farmed fish such as guppy, zebrafish, molly, swordtail,
CC   etc., in dough making and baking, in dietary aids for animals. The method
CC   provides easy manufacture of the active ingredient loaded biocompatible
CC   composition, higher yields and loading efficiency. The phytase

```
CC   incorporated in to the dietary aid is safe for animals. The present
CC   sequence is Kangaroo rat Escherichia coli phytase mutant DNA. Note: The
CC   present sequence is not shown in the specification but has been derived
CC   from Escherichia coli B phytase DNA AD050291.
XX
SQ   Sequence 1323 BP; 325 A; 353 C; 356 G; 288 T; 0 U; 1 Other;

Query Match     99.5%;  Score 1315.6;  DB 12;  Length 1323;
Best Local Similarity   99.7%;  Pred. No. 0;
Matches 1319;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 ATGAAAAGCGGATCTTAAATCCCATTTTTATCTTTCTTCTGATTCGTTAACCCCGCAAATCTGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAAAGCGGATCTTAAATCCCATTTTTATCTTTCTTCTGATTCGTTAACCCCGCAAATCTGCA 60

Qy         61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGCTGGAAAGTGTGGTGGATTCGGTTAACCGGTCATGGT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGCTGGAAAGTGTGGTGGATTCGGTTAACCGGTCATGGT 120

Qy        121 GTGCGGTGCTCCCACCAAAGCCCACCAACTGGATGCAGGAATGCACCCCAGAGGCATGGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTGCGGTGCTCCAACCAAAGCCCACCAACTGGATGCAGGAATGCACCCCAGAGGCATGGCCA 180
```

(Remaining alignment rows Qy/Db continue in the same format through position 1323, ending:)

```
Qy       1321 TAA 1323
              |||
Db       1321 TAA 1323
```

PR 24-MAY-2001; 2001US-00866379.
XX
PA (SHOR/) SHORT J M.
PA (KRET/) KRETZ K.
PA (GRAY/) GRAY K A.
PA (BART/) BARTON N R.
PA (GARR/) GARRETT J B.
PA (ODON/) O'DONOGHUE E.
PA (MATH/) MATHER E J.
XX
PI Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI Mather EJ;
XX
XX
PT Producing phytase, involves providing nucleic acid derived from bacteria
PT encoding polypeptide having phytase activity, and expressing nucleic acid
PT in yeast.
XX
DR WPI; 2004-374952/35.
DR P-PSDB; AD050292, AD050304.
XX
PS Disclosure; Page; 74pp; English.
XX
CC The invention relates to a method for producing phytase that involves
CC providing a nucleic acid encoding phytase derived from a bacteria, and
CC expressing the nucleic acid in a yeast under conditions that allow
CC expression of the enzyme in the yeast. The invention also relates to
CC modified phytase enzyme which has improved thermal tolerance and protease
CC stability at low pH. The phytase enzyme can be used in foodstuffs to
CC improve the feeding value of phytate rich ingredients, and in diet of
CC numerous animals including mammals, fowls and fishes, commercially
CC significant mammals such as pigs, goats, laboratory rodents, commercially
CC significant avian species such as chicken, ducks, doves, parrot, etc.,
CC commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC etc., in dough making and baking, in dietary aids for animals. The method
CC provides easy manufacture of the active ingredient loaded biocompatible
CC composition, higher yields and loading efficiency. The phytase
CC incorporated in to the dietary aid is safe for animals. The present
CC sequence is Escherichia coli W phytase mutant DNA. Note: The present
CC sequence is not shown in the specification but has been derived from
CC Escherichia coli B phytase DNA AD050291.
SQ Sequence 1323 BP; 324 A; 353 C; 356 G; 289 T; 0 U; 1 Other;

Query Match 99.5%; Score 1315.6; DB 12; Length 1323;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1319; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy      1 ATGAAAGGGATCTTAATCCCATTTTTATCTCTTCTTCTGATTCGTTAACCCCGCAAATCTGCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAGGGATCTTAATCCCATTTTTATCTCTTCTTCTGATTCGTTAACCCCGCAAATCTGCA 60

Qy     61 TTCGCTCAGAGTGAGCCCGAGCTGAAGCTGGAAAAGTGTGGTGATTGTCAGTCGTCATGGGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCGCTCAGAGTGAGCCCGAGCTGAAGCTGGAAAAGTGTGGTGATTGTCAGTCGTCATGGGT 120

Qy    121 GTGCGTGCTCCAACCAAGGCCACGCACGCAACTGATCATCCAGATGTCACCCCCAGACCGGCCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGCGTGCTCCAACCAAGGCCACGCACGCAACTGATCATCCAGATGTCACCCCCAGACGGCCA 180

Qy    181 ACCTGGCGCCGGTAAAACTGGGGTTGGCTGACACCCCGCCGGCTAATCGCCTATCTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACCTGGCGCCGGTAAAACTGGGGTTGGCTGACACCCCGCCGGCTAATCGCCTATCTC 240

Qy    241 GGACATTACCAACGCCAGCGTCTCGTGGTAGCCGACCGGATTGCTCGGCGAAAAAAGGGCTGCCCG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGACATTACCAACGCCAGCGTCTGGTAGCCGACCGGATTGCTCGGCGAAAAAAGGGCTGCCCG 300

Qy    301 CAGTCTCGGTCAGGTCGGCGAATTATTGCTGAATGTCGACGGAGCTGTCCGCTAAAACAGGGCGAA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGTCTCGGTCAGGTCGGCGAATTATTGCTGAATGTCGACGGAGCTGTCCGCTAAAACAGGGCGAA 360

Qy    361 GCCTTCGCGCCGCCGGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGAGATACG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCTTCGCGCCGCCGGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGAGATACG 420

Qy    421 TCCAGTCCCGATCCGTTATTTAATCCTTCTAAAAAACTGGGCGTTTTGCCAACTGGATAACGCG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCCAGTCCCGATCCGTTATTTAATCCTTCTAAAAAACTGGGCGTTTTGCCAACTGGATAACGCG 480

Qy    481 AACGGGACTGACCGGATCCTTCAGCCAGGGGCAGAGAGGGGTCAATTGCTGACTTTACCGGGGCAT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AACGGGACTGACCGGATCCTTCAGCCAGGGGCAGAGAGGGGTCAATTGCTGACTTTACCGGGGCAT 540

Qy    541 AAGGTGAGCGGCCCGACAATGTCTCATTAACCGGTGCGGTGAAGCCTTCGCGATCAATGCTGACG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAGGTGAGCGGCCCGACAATGTCTCATTAACCGGTGCGGTGAAGCCTTCGCGATCAATGCTGACG 600

Qy    601 CTTAAACGTGAGAAACAGGACGGAAAGCTGTTCATTAACGCAGGCATTAACCATCGGAACTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTTAAACGTGAGAAACAGGACGGAAAGCTGTTCATTAACGCAGGCATTAACCATCGGAACTC 660

Qy    661 GAGATAATTTCTCTGCAACAAGCACAGGAAATGCCGAGCCGGGGTGGGGAGGATCACC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GAGATAATTTCTCTGCAACAAGCACAGGAAATGCCGAGCCGGGGTGGGGAGGATCACC 720

Qy    721 GATTCACCACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGGCAATTTTATTTGCTACAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GATTCACCACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGGCAATTTTATTTGCTACAA 780

Qy    781 CGCACGCAGGAGGTTGCCCGCCAGCCCGCCACCCGTTATTGGATTGGATCATGGCACG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CGCACGCAGGAGGTTGCCCGCCAGCCCGCCACCCGTTATTGGATTGGATCATGGCACG 840

Qy    841 TTGACGCCCCATCCACCGCAAAAAACAGCGCGTATGGTGTGACATTACCCACTTCAGTACTG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TTGACGCCCCATCCACCGCAAAAAACAGCGCGTATGGTGTGACATTACCCACTTCAGTACTG 900

Qy    901 TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCCACTGGGAGCTCAACTGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCCACTGGGAGCTCAACTGG 960

Qy   1021 ACCCTTCCCGGTGCAGCCCGGATAAACACCGCCCGAGGGTGAACTGGTGAACGGCTGG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGGCTTCCCGGTCACGCCCGGATAAACACCGCCCGAGGGTGAACTGGTGAACGGCTGG 1080

Qy   1081 CGTCGGCTAAGGCGATAACGGAGGGATTCAGGTTTCGGCGTGCTGTCTTCCAGACTTTACAG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CGTCGGCTAAGGCGATAACGGAGGGATTCAGGTTTCGGCGTGCTGTCTTCCAGACTTTACAG 1140

Qy   1141 CAGATGCGTGATAAAACCCGCTGTCATTAAATACGCGCCCCGAGAGGTGAAACTGACC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAGATGCGTGATAAAACCCGCTGTCATTAAATACGCGCCCCGAGAGGTGAAACTGACC 1200

Qy   1201 CTGGCAGATGTGAAGACGGAAATGCCGCAGGGCATGTGTTCGTTGGCAGGTTTTTACGGCAA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGCAGATGTGAAGACGGAAATGCCGCAGGGCATGTGTTCGTTGGCAGGTTTTTACGGCAA 1260

Qy   1261 ATCGTGAATGAAGCCACGCATACCGGCGTGCAGTTTGAGATCTCATCACCATCAC 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAGCCACGCATACCGGCGTGCAGTTTGAGATCTCATCACCATCAC 1320

Qy   1321 TAA 1323
          |||
Db   1321 TAA 1323
```

RESULT 9
ID AAD25463
XX
AC AAD25463;
XX
ID AAD25463 standard; DNA; 1901 BP.

```
DT   26-MAR-2002  (first entry)
XX
DE   Escherichia coli appA phytase wild type DNA.
XX
KW   Bacterial phytase; K12 appA phytase; protease stability; anabolic;
KW   gastrointestinal; nutritional value; feed treatment process; therapy;
KW   thermal tolerance; growth performance; alcoholic drink; biopulping;
KW   non-alcoholic drink; biobleaching; ds.
XX
OS   Escherichia coli.
XX
FH   Key             Location/Qualifiers
FT   CDS             188..1486
FT                   /*tag= a
FT                   /product= "E. coli appA phytase protein"
XX
PN   WO20190333-A2.
XX
PD   29-NOV-2001.
XX
PF   24-MAY-2001; 2001WO-US017118.
XX
PR   25-MAY-2000; 2000US-00580515.
XX
PA   (DIVE-) DIVERSA CORP.
XX
PI   Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, O'donoghue E;
XX
DR   WPI; 2002-083108/11.
DR   P-PSDB; AAE15807.
XX
PS   Claim 42; Fig 7; 170pp; English.
XX
CC   New bacterial phytase for e.g. improving the nutritional value of phytate
CC   -containing foodstuffs and subsequently improving the growth performance
CC   of an organism that consumes it, or in treating animal digestive systems.
XX
CC   The patent discloses recombinant bacterial phytase from Escherichia coli
CC   K12 appA phytase. The enzyme has phytase activity and improved thermal
CC   tolerance when compared with wild-type phytase. It has improved protease
CC   stability at low pH. The recombinant phytase is useful for improving the
CC   nutritional value of phytate-containing foodstuffs and subsequently
CC   improving the growth performance of an organism that consumes it, in
CC   treating animal digestive systems, in feed treatment processes and for in
CC   vitro purposes related to research, discovery and development. They are
CC   also used for generating recombinant digestive system life forms, for
CC   producing or manufacturing alcoholic drinks and non-alcoholic drinks based on
CC   the use of moulds, grains and/or plants, in biopulping and bio-bleaching
CC   where a reduction in the use of environmentally harmful chemicals that
CC   are traditionally used in the pulp and paper industry is desired and in
CC   the reduction or possible elimination of the need for mineral
CC   supplements, enzymes or therapeutic drugs for animals from the daily feed
CC   thus increasing the amount calories and nutrients present in the feed.
CC   The present sequence is a DNA encoding E. coli appA phytase wild type
CC   protein.
XX
SQ   Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;

Query Match             97.4%;  Score 1287.8;  DB 6;  Length 1901;
Best Local Similarity   99.5%;  Pred. No. 0;
Matches 1293;  Conservative  0;  Mismatches  7;  Indels  0;  Gaps  0;

Qy       1 ATGAAAACGGATCTTAATCCCATTTTTATCTCTTCTTCCGTTAACCCCGCAAATCTGCA 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     188 ATGAAAACGGATCTTAATCCCATTTTTATCTCTTCTTCCGTTAACCCCGCAAATCTGCA 247

Qy      61 TTCGCTCAGAGTGAGTGAGCCGGAGCTGAAAGTGTGGTGGTGATTGTCAGTCGTCATGGT 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     248 TTCGCTCAGAGTGAGTGAGCCGGAGCTGAAAGTGTGGTGGTGATTGTCAGTCGTCATGGT 307

Qy     121 GTGCGTGCTCCAACCAAGGACGCAACTGATGCAGGATGTCACCCCAGACGCCATGGCCA 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     308 GTGCGTGCTCCAACCAAGGACGCAACTGATGCAGGATGTCACCCCAGACGCCATGGCCA 367

Qy     181 ACCTGGCCGGTAAAACTGGGGTTGGCCTGACACCGCGGGNGGTGGTGAGCTAATCGCTC 240
           |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     368 ACCTGGCCGGTAAAACTGGGGTTGGCCTGACACCGCGGGGGGTGGTGAGCTAATCGCCTATCTC 427

Qy     241 GGACATTACCAAACGCCAGGTCTTCGGTAGCCGCCGATTTGCTGGGCGGAAAAAGGGCTGCCCG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     428 GGACATTACCAAACGCCAGGTCTTCGGTAGCCGCCGATTTGCTGGGCGGAAAAAGGGCTGCCCG 487

Qy     301 CAGTTCGTCAGTTCGGCGATTATTGCTGATTGTGAGTGTCGACGAGCGGTACCCGTAAAACAGGCGAA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     488 CAGTTCGTCAGTTCGGCGATTATTGCTGATTGTGAGTGTCGACGAGCGGTACCCGTAAAACAGGCGAA 547

Qy     361 GCCTTTCGCGCGCCGCCGGGCTGACACTGTCGCAATAAACCAGGCGAGATACG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     548 GCCTTTCGCGCGCCGCCGGGCTGACACTGTCGCAATAAACCAGGCGAGATACG 607

Qy     421 TCCAGTCCCGATCCGTTATTTAATCCCTCTAAAAAACTGGGGTTTGCCAACTGGATAACGCG 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     608 TCCAGTCCCGATCCGTTATTTAATCCCTCTAAAAAACTGGGGTTTGCCAACTGGATAACGCG 667

Qy     481 AACGTGACTGAGACGGCCGGATCCTCAGCAGGGCAGGAGGGGTCAAATTGCTCAATTGCTGACTTTACCGGGCAT 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     668 AACGTGACTGAGACGGCCGGATCCTCAGCAGGGCAGGAGGGGTCAAATTGCTCAATTGCTGACTTTACCGGGCAT 727

Qy     541 CGGCAAAACCGGCGTTTCGGGAACTGGAACGGGGGTGCTTAATTTTCCGCAATCAAAACTTGTGC 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     728 CGGCAAAACCGGCGTTTCGGGAACTGGAACGGGGGTGCTTAATTTTCCGCAATCAAAACTTGTGC 787

Qy     601 CTTAAACGTTGAGAAAACAGGACGAAAGCTGTTCAATTAACCAGGCATTACCCATCGGAACTC 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     788 CTTAAACGTTGAGAAAACAGGACGAAAGCTGTTCAATTAACCAGGCATTACCCATCGGAACTC 847

Qy     661 AAGGTGAGCGGCCGACAATGTCTTCATTAACCGGTGCGGTRAAGCCTGCAATCAAATGCTGACG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     848 AAGGTGAGCGGCCGACAATGTCTTCATTAACCGGTGCGGTAAGCCTGCAATCAAATGCTGACG 907

Qy     721 GAGATATTTCTCCTGCAACAAGACACAGGGAAGTGCCGGACCGGGGAAGGAATCACC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     908 GAGATATTTCTCCTGCAACAAGACACAGGGAAGTGCCGGACCGGGGAAGGAATCACC 967

Qy     781 GAATTCACACCAGTGGGAACACCTTGCTCAAGTTTGCGATAACGCGCAAATTTTATTTGCTACAA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     968 GAATTCACACCAGTGGGAACACCTTGCTCAAGTTTGCGATAACGCGCAAATTTTATTTGCTACAA 1027

Qy     841 CGCACGCCAGAGGTTGCCCGCGCGCCAGCCGCCCCCGTTATTGGAATTGGATTGATCATGGCAGCG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1028 CGCACGCCAGAGGTTGCCCGCGCGCCAGCCGCCCCCGTTATTGGAATTGGATTGATCAGACAGCG 1087

Qy     901 TTGACGCCCCATCCACCCGCAAAAACAGGCGTATGGTGGTGGTGACATTACCCACTCAGTACTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1088 TTGACGCCCCATCCACCCGCAAAAACAGGCGTATGGTGGTGGTGACATTACCCACTCAGTGCTG 1147

Qy     961 TTTATTGCCGGACACGGATACTAATCTGGCAAAAATCTCGCGCGGCAAAACTCGAGCTCAAACTGG 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1148 TTTATTGCCGGACACGGATACTAATCTGGCAAAAATCTCGCGCGGCAAAACTCGAGCTCAAACTGG 1207

Qy    1021 ACGGCTTCCCGGTCAGCCGGATAACACGCGCCAGGTGGTGAACTGGTGTTTGAACGCGCTGG 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1208 ACGGCTTCCCGGTCAGCCGGATAACACGCGCCAGGTGGTGAACTGGTGTTTGAACGCGCTGG 1267

Qy    1081 CGTCGGCTAAGCCGATAACAGCCAGTGGATTCAGGTTTCGGTCTGGTCTTCCAGACTTTACAG 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1268 CGTCGGCTAAGCCGATAACAGCCAGTGGATTCAGGTTTCGGTCTGGTCTTCCAGACTTTACAG 1327

Qy    1141 CAGATGCGTTGAATAAAACCCGCTGTCAGTCATTAAATTACGGCGCCCGGAGAGGTGAAAACTGACC 1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1328 CAGATGCGTTGAATAAAACCCGCTGTCAGTCATTAAATTACGGCGCCCGGAGAGGTGAAAACTGACC 1387

Qy    1201 CTGGCAGGAGGTGAAGAGCGGAAAGCGGAAAGCGGCAGTGTTCGTTGGGCAGGTTTTACGGCAA 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1388 CTGGCAGGAGGTGAAGAGCGGAAAGCGGAAAGCGGCCATGTGTTCGTTGGGCAGGTTTTACGGCA 1447
```

Qy   1261 ATCGTGAATGAAGGACAGGCATACCGGGGCGTGCAGTTTGAGAT 1300
           |||||||||||||||||||||||||||||||||||||||||||
Db   1448 ATCGTGAATGAAGGCACCGCATACCGGGCGTGCAGTTTGTAAT 1487

RESULT 10

ADA19449
ID   ADA19449 standard; DNA; 1901 BP.
XX
AC   ADA19449;
XX
DT   20-NOV-2003 (first entry)
XX
DE   E. coli K12 appA gene encoding phytase.
XX
KW   Phytase; de; gene; phytate; appA gene; animal feed; inorganic phosphate;
KW   digestion enhancement; transgenic; thermal tolerance; protease stability.
XX
OS   Escherichia coli; strain K12.
XX
FH   Key   Location/Qualifiers
FT   CDS   188..1486
FT   /*tag= a
FT   /product= "Phytase"
XX
PN   US2002136754-A1.
XX
PD   26-SEP-2002.
XX
PF   24-MAY-2001; 2001US-00866379.
XX
PR   13-AUG-1997; 97US-00910798.
PR   01-MAR-1999; 99US-00259214.
PR   13-APR-1999; 99US-00291931.
PR   25-MAY-1999; 99US-00318528.
PR   25-MAY-2000; 2000US-00580515.
XX
PA   (SHOR/) SHORT J M.
PA   (KRET/) KRETZ K A.
PA   (GRAY/) GRAY K A.
PA   (GRAY/) GRAY K A.
PA   (BART/) BARTON N R.
PA   (GARR/) GARRETT J B.
PA   (DONO/) O' DONOGHUE E.
PA   (MATH/) MATHUR E J.
XX
PI   Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB;
PI   O' Donoghue E, Mathur EJ;
XX
DR   WPI; 2003-040002/03.
DR   P-PSDB; ADA19450.
XX
PT   Isolated Escherichia coli polynucleotide encoding a modified phytase
PT   enzyme, useful in the production of animal feed, for improving the
PT   nutritional value of phytate-containing foodstuff and for enhancing
PT   digestion in humans and animals.
XX
PS   Claim 42; Fig 7; 62pp; English.
XX

The invention relates to an isolated Escherichia coli polynucleotide
encoding a phytase enzyme appearing as ADA19450 and having amino acids
modification from W68E, Q84W, A95P, K97C, S168B, R180Y, N226C or Y277D.
Also included the E. coli appA gene ADA19449 (or an oligonucleotide
derived from it) or its mutant sequence ADA19452, expression vectors,
host cells, a method of improving nutritional value of a phytate-
containing foodstuff by contacting the phytate-containing foodstuff with
a pure phytase enzyme mutated phytase (where the phytase enzyme catalyzes
the liberation of inorganic phosphate from the phytate in the phytate-
containing foodstuff), a method to produce an animal feed containing a
microbial phytase (comprising culturing the plant cell, plant part or
plant under conditions where the nucleotide sequence is expressed and
converting the plant cells, plants or plant into a composition for animal
feed), a feed composition for animals (comprising the plant seeds, plant
cells, plant parts or plants in admixture with a phytate-containing

CC   foodstuff), a method to treat a human or an animal able to benefit from
CC   digestive enhancement by the activity of an exogenous phytase enzymes
CC   comprising administering to the human or animal the plant seed, plant
CC   cells, plant parts or plants of a transgenic plant which expresses a
CC   contain an expression system which expresses a nucleotide sequence
CC   encoding a phytase enzyme, a transgenic non-human organism whose genome
CC   comprising a heterologous nucleic acid sequence encoding a polypeptide
CC   having phytase activity. The phytase enzyme is useful for improving the
CC   nutritional value of phytate-containing foodstuff, in the production of
CC   animal feed and for enhancing digestion in humans and animals. The
CC   invented method improves thermal tolerance and protease stability. It
CC   also improves the feeding value of phytate rich ingredients. The present
CC   sequence represents the E. coli K12 appA gene encoding wild-type phytase.
XX
SQ   Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;

Query Match   97.4%; Score 1287.8; DB 8; Length 1901;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy   1 ATGAAAGCGGATCTTAATCCCATTTTTAATCTCTTCTGATTCCGTTAACCCCGCAAATCTGCA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 ATGAAAGCGGATCTTAATCCCATTTTTAATCTCTTCTGATTCCGTTAACCCCGCAAATCTGCA 247

Qy   61 TTCGCTCAGAGTGAGCCGGGAGCGTGAAGCTTGGAAAGTGTGGGGATTGTCAGTCGTCATGGT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248 TTCGCTCAGAGTGAGCCGGAGCGTGAAGCTGGAAAGTGTGGGGATTGTCAGTCGTCATGGT 307

Qy   121 GTGCGTGCCTCCAACCAAGGCCACGCAACTGATGGCAGCGGATGTCACCCCCAGACGCCATGGCCA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308 GTGCGTGCTCCAACCAAGGGCCACGCAACTGATGGCAGGATGTCACCCCCAGACGCCATGGCCA 367

Qy   181 ACCTGGGCCGGTAAAACTGGGTTGGCTGACACCGCGGNGGTGGTGAGCTAATCGGCCTATCTC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368 ACCTGGGCCGGTAAAACTGGGTTGGCTGACACCGCGGNGGTGGTGAGCTAATCGGCCTATCTC 427

Qy   241 GGACATTACCAACGCACGGCCGTCTGGGTAGCCGACGGACGGATTGCTGGCGGAAAAAAGGCTCGCCG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 GGACATTACCAACGCACGGCCGTCTGGGTAGCCGACGGACGGATTGCTGGCGGAAAAAGGCTCGCCG 487

Qy   301 CAGTGCTCGGTCAGGTCGCGGCGGATTGATGTCGACCGTACCCCTAAAACAGCGGAA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488 CAGTGCTCGGTCAGGTCGCGGCGGATTGATGTCGACCGTACCCCTAAAACAGCGGAA 547

Qy   361 GCCTTCGCGCGCCGGGCTGGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548 GCCTTCGCGCGCCGGGCTGGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 607

Qy   421 TCCAGTCCCGAACCCGGTTATTTAATCCTCTAAAAAACTGGGCGTTTGCCAAACTGGATAACGCG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   608 TCCAGTCCCGATCCGTTATTTTAATCCTCTAAAAAACTGGGCGTTTGCCAAACTGGATAACGCG 667

Qy   481 AACGTGACTGACGGCGGATCCTCAGCGACGGCAGAGGGGTCAAATGCTGACTTTACCGGGCAT 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668 AACGTGACTGACGGCGATCCTCAGCGACGGCAGAGGGGTCAAATGCTGACTTTACCGGGCAT 727

Qy   541 CGGCAAACGGCGGTTTCGGGAACTGGGAACGGGTGGCTTAATTTTCCGGGCAATCAAACTTGTGC 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   728 CGGCAAACGGCGGTTTCGGGAACTGGGAACGGGTGGCTTAATTTTCCGGGCAATCAAACTTGTGC 787

Qy   601 CTTAAACGTGAGAAAACAGGAGGACGGAAAGGTGTTCATTAACGGCAGGCATTACCATCGGAACTC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   788 CTTAAACGTGAGAAAACAGGACGGAAAGGTGTTCATTAACGGCAGGCATTACCATCGGAACTC 847

Qy   661 AAGGTGAGCGGCGCGACAAATGTCTCAATTAACCGGTGGGCGTTAAGCCTCGGCATCAATGCTGACG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   848 AAGGTGAGCGGCGCGACAAATGTCTCAATTAACCGGTGGGCGTTAAGCCTCGGCATCAATGCTGACG 907

Qy   721 GAGATATTTCTCCTGCAACAACGCACAGGGAATGCCGGGCGGCCGGGGCCGGGGGGAGGATACC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   908 GAGATATTTCTCCTGCAACAACGCACAGGGAATGCCGGAGCGCGGGGGGCGCGGGGAAGGAATCACC 967

Qy   781 GAATTCACCACGGAACACCTTGCTAAGTTTGCATAACGGCAAATTTTATTTGCTACAA 840

```
RESULT 11
ADC87744
ID   ADC87744 standard; DNA; 1901 BP.
XX
AC   ADC87744;
XX
DT   01-JAN-2004  (first entry)
XX
DE   DNA encoding Escherichia coli appA phytase.
XX
KW   Phytase; food supplement; enzyme delivery matrix; soybean meal;
KW   thermotolerance; thermostability; kernel; phytate;
KW   myo-inositol-hexaphosphate; inositol; inorganic phosphate;
KW   thermotolerant; feed value; digestion; gene; ds; appA.
OS   Escherichia coli.
XX
FH   Key             Location/Qualifiers
FT   CDS             188..1486
FT                   /*tag= a
FT                   /product= "AppA phytase"
FT                   /transl_except= (pos:401..403,aa:Arg)
XX
PN   US2003103958-A1.
XX
PD   05-JUN-2003.
XX
PF   24-MAY-2002; 2002US-00156660.
XX
PR   13-AUG-1997; 97US-00910798.
PR   01-MAR-1999; 99US-00259214.
PR   13-APR-1999; 99US-00291931.
PR   25-MAY-1999; 99US-00318528.
PR   25-MAY-2000; 2000US-00580515.
PR   24-MAY-2001; 2001US-00866379.
XX
PA   (DIVE-) DIVERSA CORP.
```

```
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI   Mathur EJ;
XX
DR   WPI; 2003-787039/74.
DR   P-PSDB; ADC87745.
XX
PT   New nucleic acid encoding a polypeptide having phytase activity, useful
PT   in improving the feeding value of phytate rich ingredients or as an aid
PT   in phytate digestion.
XX
PS   Example 1; SEQ ID NO 3; 1113pp; English.
XX
CC   The invention discloses a new isolated or recombinant nucleic acid which
CC   encodes a polypeptide having a phytase activity. Also claimed is a
CC   nucleic acid probe, an amplification primer sequence pair, an expression
CC   cassette comprising the nucleic acid, a vector comprising the nucleic
CC   acid, a transgenic non-human animal or plant, or its seed, comprising the
CC   nucleic acid, an antisense oligonucleotide, inhibiting the translation of
CC   a phytase message in a cell, a heterodimer comprising the polypeptide and
CC   a second domain, an array comprising immobilised polypeptide or nucleic
CC   acid, a hybridoma comprising an antibody that specifically binds to the
CC   polypeptide, a food supplement for an animal, an edible delivery
CC   matrix, an edible pellet comprising a granule edible carrier and the
CC   polypeptide, a feed composition, a soybean meal, isolating or identifying
CC   the polypeptide, making an anti-phytase antibody, producing a recombinant
CC   polypeptide, determining whether a compound binds to the polypeptide,
CC   identifying a modulator, whole cell engineering of new or modified
CC   phenotypes by using real-time metabolic flux analysis, increasing
CC   thermotolerance or thermostability of the phytase polypeptide, increasing
CC   the resistance of the phytase polypeptide to enzymatic inactivation in a
CC   digestive system and processing of corn and sorghum kernels. The phytase
CC   activity comprises catalysis of phytate (myo-inositol-hexaphosphate) to
CC   inositol and inorganic phosphate and is thermotolerant. The nucleic acid
CC   is useful in improving the feeding value of phytate rich ingredients or
CC   as an aid in phytate digestion. The sequence presented is the DNA
CC   encoding the Escherichia coli appA phytase.
XX
SQ   Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;
```

```
Query Match           97.4%;  Score 1287.8;  DB 10;  Length 1901;
Best Local Similarity 99.5%;  Pred. No. 0;
Matches 1293; Conservative  0; Mismatches  7; Indels  0;  Gaps  0;
```

```
Qy        1 ATGAAAGCGATCTTAATCCCATTTTTATCTTCTTCTGATTCCGTTAAACCCGCAAATCTGCA 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      188 ATGAAAGCGATCTTAATCCCATTTTTATCTTCTTCTGATTCCGTTAAACCCGCAAATCTGCA 247

Qy       61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      248 TTCGCTCAGAGTGAGCCGGAGCTGAAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 307

Qy      121 GTGGGTGCTCCACCAAGCCCAGGCACTGATGCAGGAGTCACCCCACGACATGGCCA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      308 GTGGGTGCTCCAACCAAGGCCACCGGACGGATTGCCTGGGGAAAAAGGGCTGCCCG 367

Qy      181 ACCTGGGCCCGGTAAAAACTGGGTTGGCTGCAGACCCGCGGNGTGGGGTGAGCTAATCGCCTATCTC 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      368 ACCTGGGCCCGGTAAAAACTGGGTTGGCTGACACCCGCGGNGTGGGGTGAGCTAATCGCCTATCTC 427

Qy      241 GGACATTACCAACGCCAGCGGTCTGGTAGCCGACGGATTGCTGGGGCGAAAAAGGGGCCA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      428 GGACATTACCAACGCCAGCGGTCTGGTAGCCGACGGATTGCTGGGGCGAAAAAGGGCTGCCCG 487

Qy      301 CAGTCTCGTGTCAGTCGCGCATTATTGCTGATGTGTCGCAATAACCGGTACACTACCCAGGAGATACG 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      488 CAGTCTGGTCAGGTCGCCGGATTATTGCTGATGTGTCGACGAACCGTACCCGTAAAAACAGGCCGAA 547

Qy      361 GCCTTCGCCGCCGGCGGCTCGGCACCTGACTGTGCAATAACCGGTACACTACCCAGGAGATACG 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      548 GCCTTCGCCGCCGGCGGCTCGGCACCTGACTGTGCAATAACCGGTACACTACCCAGGAGATACG 607

Qy      421 TCCAGTCCGGATCCCGTTATTAATCCTCTAAAAAACTGGGGTTTGCCAACTGGATAACGCG 480
```

```
Db      968 GATTCACCACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA 1027
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      841 CGCACGCCAGAGAGTTGCCCGCGACGCCGCCCCCGTTATTGATTTGATTCATGGCCACGG 900

Qy      901 TTGACGCGCCCATCCACCCGCAAAAAAGGGGTATGGTGTGACATTACCCACTTCAGTACTG 960
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1028 CGCACGCCAGAGGTTGCCCGCGACGCCCACCCCGTTATTAGATTTTGATTCAAGACACGG 1087

Qy      961 TTTATTGCCGGACACGATACTAATCTGGCAAAATCTCGGCGCGGCGCACTGAGCTCAACTGG 1020
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1088 TTGACGCCCATCCACCCGCAAAAAACAGGCGTATAGGTGTGACATTACCCACTTCAGTGCTG 1147

Qy     1021 ACGGCTTCCCGGTCAGCCGGATAACACGCGCGCCCAGGGTGCTGAAACTGGGTCGTGTTGAACGCTGG 1080
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1148 TTTATTGCCGGACACGATACTAATCTGGCAAAATCTCGGCGCGGCGCACTGAGCTCAACTGG 1207

Qy     1081 CGTCGGCTAAGCGGATAACAGCCAGTGGAATTCAGGTTGCGCTGGTGTCTTCCAGACTTTACAG 1140
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1208 ACGGCTTCCCGGTCAGCCGGATAACACGCGCGCCCAGGGTGCTGAAACTGGGTCGTGTTGAACGCTGG 1267

Qy     1141 CAGATGCCTGATAAAAACCGCGCTGTCATTAAAATACGCCCCGGAGGAGGAGTGAAAACTGACC 1200
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1268 CGTCGGCTAAGCGGATAACAGCCAGTGGAATTCAGGTTGCGCTGGTGTCTTCCAGACTTTACAG 1327

Qy     1201 CTTGGCCAGGATGTGAAAGAGGGAAAATGCCGAGGGCCATGCGTGTTCGTTCGTTACGGAA 1260
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1328 CAGATGCCTGATAAAAACCGCGCTGTCATTAAAATACGCCCCGGAGGAGGAGTGAAAACTGACC 1387

Qy     1261 ATCGTGAATGAAGGCACGCATACCGGCGTGCAGTTTGAGAT 1300
             ||||||||||||||||||||||||||||||||||||| ||
Db     1388 CTTGGCCAGGATGTGAAAGAGGGAAAATGCCGAGGGCCATGCGTGTTCGTTCGTTACGGAA 1447

Qy     1448 ATCGTGAATGAAGGCACGCATACCGGCGTGCAGTTTGTAAT 1487
             ||||||||||||||||||||||||||||||||||||| ||
Db     1448 ATCGTGAATGAAGGCACGCATACCGGCGTGCAGTTTGTAAT 1487
```

```
RESULT 12
ID   ADO50297 standard; DNA; 1901 BP.
XX
AC   ADO50297;
XX
DT   29-JUL-2004 (first entry)
XX
DE   Escherichia coli K12 appA phytase DNA.
XX
KW   AppA phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW   phytate; animal feed; fish feed; dough; baking; gene; ds.
XX
OS   Escherichia coli; K12.
XX
FH   Key             Location/Qualifiers
FH   CDS             188..1486
FT                   /*tag= a
FT                   /product= "appA phytase"
FT   unsure          401..403
FT                   /*tag= b
FT                   /note= "Encodes Arg"
XX
PN   US20040091968-A1.
XX
PD   13-MAY-2004.
XX
PF   20-JUN-2003; 2003US-00601319.
XX
PR   13-AUG-1997; 97US-00910798.
PR   01-MAR-1999; 99US-00259214.
PR   13-APR-1999; 99US-00291931.
PR   25-MAY-1999; 99US-00318528.
PR   25-MAY-2000; 2000US-00580515.
PR   24-MAY-2001; 2001US-00866379.
XX
PI   SHORT J M, KRETZ K, GRAY KA, BARTON NR, GARRETT JB, O'donoghue E;
PI   Mather EJ;
XX
PA   (SHOR/) SHORT J M.
PA   (KRET/) KRETZ K.
PA   (GRAY/) GRAY K A.
PA   (BART/) BARTON N R.
PA   (GARR/) GARRETT J B.
PA   (ODON/) O'DONOGHUE E.
PA   (MATH/) MATHER E J.
XX
PT   Producing phytase, involves providing nucleic acid derived from bacteria
PT   encoding polypeptide having phytase activity, and expressing nucleic acid
PT   in yeast.
XX
PS   Disclosure; SEQ ID NO 7; 74pp; English.
XX
DR   WPI; 2004-374952/35.
DR   P-PSDB; ADO50298.
XX
CC   The invention relates to a method for producing phytase that involves
CC   providing a nucleic acid encoding phytase derived from a bacteria, and
CC   expressing the nucleic acid in a yeast under conditions that allow
CC   expression of the enzyme which has improved thermal tolerance and protease
CC   modified phytase enzyme in the yeast. The invention also relates to
CC   stability at low pH. The phytase enzyme can be used in foodstuffs to
CC   improve the feeding value of phytate rich ingredients, and in diet of
CC   numerous animals including mammals, fowls and fishes, commercially
CC   significant mammals such as pigs, goats, laboratory rodents, commercially
CC   significant avian species such as chicken, ducks, doves, parrot, etc.,
CC   commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC   etc., in dough making and baking, in dietary aids for animals. The method
CC   provides easy manufacture of the active ingredient loaded biocompatible
CC   composition, higher yields and loading efficiency. The phytase
CC   incorporated in to the dietary aid is safe for animals. The present
CC   sequence is Escherichia coli K12 appA phytase DNA.
XX
SQ   Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;

Query Match             97.4%;  Score 1287.8;  DB 12;  Length 1901;
Best Local Similarity   99.5%;
Matches 1293;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy     1  ATGAAAACGGATCTTAATCCCCATTTTTTATCTCTTCCGTTAACCCCGCCAATCTGCA   60
          ||||||| |||||||||||||| |||||||| ||||||||||||||||||||||||||
Db   188  ATGAAAAGCGATCTTAATCCCATTTTTTATTTATCTCTTCCGTTAACCCCGCCAATCTGCA  247

Db   248  TTCGCCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGGTGATTGTCAGTCGTCATGGT  307
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    61  TTCGCCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGGTGATTGTCAGTCGTCATGGT  120
```

```
Db   608  TCCAGTCCGATCCGGTTATTTAATCCCTCTAAAAAACTGGCGTTGCCAACTGGATAACGCG  667
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   481  AACGTGACTGAGCGATCCTCAGCAGGCAGCAGGGTTCAATTGCTGACTTTACCGGGCAT   540

Db   668  AACGTGACTGAGCGATCCTCAGCAGGGCAGAGAGGGTCAATTGCTGACTTTACCGGGCAT  727
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   541  CGGCAAACGGCGTTCGGCAACTGGGAACGGGTGCTTAATTTCCGCAATCAAAACTTGTGC  600

Db   728  CGGCAAACGGCGTTCGGCAACTGGGAACGGGTGCTTAATTTCCGCAATCAAAACTTGTGC  787
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   601  CTTAAACGTGAGAAAACAGGAACGAAAGCTGTTCATTAACGCAGGCCATTACCATCGGAACTC  660

Db   788  CTTAAACCTGAGAAAACAGGAACGAAAGCTGTTCATTAACGCAGGCCATTACCATCGGAACTC  847
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   661  AAGGTGAGCGCGGACAAATGTCTCATTAACCGGTGCGCGTAAGCCTTCGCATCAAATGCTGACG  720

Db   848  AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTTCGCATCAAATGCTGACG  907
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   721  GAGATATTTCTCCTGCAACAAGCACCAGGGAATGCCGGAGCCGGGGGTGGGGAGGGATCACC  780

Db   908  GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCCGGAGCCGGGGGTGGGGAGGGATCACC  967

Qy   841  CGCACCCCAGAGAGGTTGCCCGCCGAGCCCGCCGCCACCCCGTTATTACGCGCAGCAGCG  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   968  GATTCACACCAGTGGGACACCCTTGCTAAGTTTGCATAAGCGCGCAAATTTTATTTGCTACAA  1027

Qy   961  TTTATTGCCGGACACGATACTAATCTCCGCGGCCCACTGGAGCTCAACTGG  1020
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1088  TTGACGCGCCCATCCACCGCAAAAAACAGGCGTATTGGTGTGACATTACCCACTTCAGTACTG  1147

Qy   901  TTGACGCGCCCATCCACCGCAAAAAACAGGCGTATTGGTGTGACATTACCCACTTCAGTACTG  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1028  CGCACGCCCAGAGGTTGCCCGCGCCAGCCGCCACCCCGTTATTGCGCCAGCAGCG  1087

Qy  1021  ACGCTTCCCGGTCAGCCGGATAACACCGCGCCGGATAAACAGCGGTGGTGAGTGTTTGAACGGCTGG  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1148  TTTATCGCCGGACACGATACTAATCTCCGCGGCCCACTGGAGCTCAACTGG  1207

Qy   781  GATTCACACCAGTGGGACACCCTTGCTAAGTTTGCATAAGCGCGCAAATTTTATTTGCTACAA  840

Db   968  GATTCACACCAGTGGGACACCCTTGCTAAGTTTGCATAAGCGCGCAAATTTTATTTGCTACAA  1027

Qy  1081  CGTCGGGTTAAGCGAATAACAGCCAGTGAGTTCAGGTTCGCTGGTCTTCCAGACTTTACAG  1140

Db  1268  CGTCGGGCTAAGCGAATAACAGCCAGTGAGTTCAGGTTCGCTCGTCTTCCAGACTTTACAG  1327

Qy  1141  CAGATGCCTGATAAAAACCGCGCTGTCATTAAATACGGCCGCCCGAGACGGTGAAAACTGACC  1200

Db  1328  CAGATGCCTGATAAAAACCGCGCTGTCATTAAATACGGCCGCCCGAGACGGTGAAAACTGACC  1387

Qy  1201  CTGGCACGGATGTGAAGGAGGCAGGCCATGTGTCGTTCGTTGGCGAGGTTTTACCGGAA  1260

Db  1388  CTGGCACGGATGTGAAGACGGAAATGCCGCAGGCCATGTGTCGTTCGTTGGCGAGGTTTTACCGGAA  1447

Qy  1261  ATCGTGAATGAAGCACGGCATACCGGGTGCAGTTTGAGAT  1300

Db  1448  ATCGTGAATGAAGCACCGGCATACCGGGCTGCAGTTTGTAAT  1487
```

```
Qy    121  GTGGGTGCTCCAACCAAGGCCACCGCAACTGGATGTCACCCCAGACGCGATGGCCA        180
           |||||||||||||||||||||||||||||||||||| |  |||||||||||||||
Db    308  GTGGGTGCTCCAACCAAGGCCACCGCAACTGATGCGCAGGACGCGATGGCCA           367

Qy    181  ACCTGGCCCGGTAAAACTGGGGTTGGCTGACACCGGCGNGGTGGTGAGCTAATCGCCTATCTC  240
           ||||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db    368  ACCTGGCCCGGTAAAACTGGGTTGGCTGACACCGGCGGNGGTGGTGAGCTAATCGCCTATCTC  427

Qy    241  GGCCATTACCAACCCAGGGTCTCGGTAGCCGACACGGATTGCTCGGGGAAAAAGGGGCTGGCCCG  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428  GGACATTACCAACGCCAGCGGTCTGGTAGCCGACGATTGCTGGGCCGAAAAAGGGGCTGGCCCG   487

Qy    301  CAGTCTGGTCAGGTCAGGTTCGCGGATTATTGCTGATAGTGTCGACGAGCGCGTAAAAACAGGCGAA  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488  CAGTCTGGTCAGGTTCGCGGATTATTGCTGANTGTCGACGAGACGGCGTACCGTAAAAACAGGCGAA  547

Qy    361  GCCTTCGCCGCCCGGGGCTGCACCTGACTGTTGCAATAACCGTACATACCAGGCAGATACG     420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548  GCCTTCGCCGCCCTGACTGTTGCAATAACCGTGCAATAACCGTACATACCAGGCAGATACG    607

Qy    421  TCCAGTCCCGATCGTTCCGTTATTTAATCCTCTAAAAAACTGGGCGTTTGCCCAACTGGATAAACCGG  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608  TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAAACTGGGCGTTTGCCAACTGGATAAACCGG     667

Qy    481  AACCTGACTGACGGCGATCCTCAGCAGGGGCGAGGGGTCAATTGCTGACTTTACCGGGGCAT      540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    668  AACGTGACTGACGGCGATCCTCAGCAGGGGCGAGGGGTCAATTGCTGACTTTACCGGGCAT     727

Qy    541  CGGCAAACCGCGTTTCGCGAACTGACCGGTGCTTAATTTCCGCAATCAAACTTGTGC        600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    728  CGGGCAAACCGCGTTTCGCGAACTGGAACCGGCTGCTTAATTTCCGCAATCAAAACTTGTGC  787

Qy    601  CTTAAACTGAGAAACACAGAGGACGAAAGCTGTTTCATTAAAGGCGGCCATTACCATCGGGAACTC   660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    788  CTTAAACGTGAGAAACAGAGGACGAAAGCTGTTCATTAAAGGCGGCCATTACCATCGGGAACTC    847

Qy    661  AAGGTGAGCGCCCGACAAGTCTCATTCATTAACCGGTGCGGGTAAGCCTTCGCATCAATGCTGAGG   720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    848  AAGGTGAGCGCCGACAAGTCTCATTCATTAACCGAGGGCATGTTAAGCCTTCGCATCAATGCTGAGG  907

Qy    721  GAGATATTTCTCCTGCACAAGCACAGGGAATGCCGGACCCGGGGTGGGAAGGATCACC      780
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908  AACGTGACTGACGGCGATCCTCAGCAGGGGCGAGGGGTCAATTGCTGACTTTACCGGGCAT     967

Qy    781  GATTCACACCAGTGGAACACCTTGCTAAGTTGCATAAACGGCGAACATTTATTTGCTACAA   840
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968  GATTCACACCAGTGGAACACCTTGCTAAGTTGCATAAACGGCGAATTTATTTGCTACAA     1027

Qy    841  CGCACGCCAGAGGTTGCCCGCAGCCGGCGCCACCCCGTATTGATTGGATTCATTGGCAGCG   900
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908  GAGATATTTCTCCTGCACAAGCACAGGGAATGCCGGACCCGGGGTGGGAAGGATCACC      967

Qy    901  TTGACGCCCATCCACCGCAAAAACAGGCGTATGGTGTTGACATTACCCACTTCAGTGCTG   960
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088  TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCTG    1147

Qy    961  TTTATTCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGGCTCAACTGG        1020
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148  TTTATTCGCCGGACACGATACTAATCTGGCAAATCTCGGAGCTCAACTGG              1207

Qy   1021  ACGCTTCCCGGTCCAGCCGGGATAACACGGCGGATAACACGGCGGCGCTTTGAAACGGCTGG   1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208  ACGCTTCCCGGTCAGCCAGGTGGTGGTGAACTGGTTTGAAACGGCTGG              1267

Qy   1081  CGTCGGCTAAGCGATAACAGTGGATTCAGTTTTCGCTCGTTCTTCCAGACTTTACAG       1140
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   1268  CGTCGGCTAAGCGATAACAGCAGTGGATTCAGGTTTCGCTGGTTCTTCCAGACTTTACAG    1327

Qy   1141  CAGATGGCTGATAAAAACGCGCTGTCATTAAATACGCGCCCGCTGTCATTAAAATACGCGGTGAAACTGACC  1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328  CAGATGCTGATAAAAACCCGCTGTCATTAAAATACGCGGTGAAACTGACC           1387
```

```
Qy   1201  CTGGCCAGGATGTGAAGAGGCAAATGCGCCAGGGCCATGTGTTCGTTCGGCTTTTACGGCAA  1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388  CTGGCCAGGATGTGAAGAGGCAAATGCGCCAGGGCCATGTGTTCGTTCGGCTTTTACGGCAA  1447

Qy   1261  ATCGTGAAATGAAGCACGCATACCGGCGCGTGCAGTTTGAGAT   1300
           ||||||||||  |||||||||||||||||||||||| |||||
Db   1448  ATCGTGAATGAAGCACGCATACCGGCGCGTGCAGTTTGTAAT    1487
```

RESULT 13

AED50819

ID   AED50819 standard; DNA; 1901 BP.
XX
AC   AED50819;
XX
DT   29-DEC-2005 (first entry)
XX
DE   Escherichia coli phytase appA gene.
XX
KW   Transgenic animal; gene therapy; feedstuff; food; thermostable; ds; gene;
KW   3-phytase; appA gene.
XX
OS   Escherichia coli; strain K12.
XX
FH   Key   Location/Qualifiers
FH
FT   CDS   188..1486
FT   /*tag= a
FT   /product= "Escherichia coli phytase protein"
FT   /transl_except= (pos:401..403,aa:Arg)
FT   /EC_number= "3.1.3.8"
XX
PN   US2005246780-A1.
XX
PD   03-NOV-2005.
XX
PF   11-FEB-2005; 2005US-00056354.
XX
PR   13-AUG-1997; 97US-00910798.
PR   01-MAR-1999; 99US-00259214.
PR   13-APR-1999; 99US-00291931.
PR   25-MAY-1999; 99US-00318528.
PR   25-MAY-2000; 2000US-00580515.
PR   24-MAY-2001; 2001US-00866379.
PR   24-MAY-2002; 2002US-00156660.
XX
PA   (SHOR/) SHORT J M.
PA   (KRET/) KRETZ K.
PA   (GRAY/) GRAY K A.
PA   (BART/) BARTON N R.
PA   (GARR/) GARRETT J B.
PA   (ODON/) ODONOGHUE E.
PA   (MATH/) MATHUR E J.
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, Odonoghue E;
PI   Mathur EJ;
XX
DR   WPI; 2005-733851/75.
DR   P-PSDB; AED50820.
DR   GENBANK; M58708.
XX
PT   New isolated or recombinant nucleic acid encoding phytase enzymes, useful
PT   as foodstuff, and for oil degumming, producing an animal feed, delivering
PT   a phytase enzyme supplement to an animal.
XX
PS   Example 1; SEQ ID NO 3; 104pp; English.
XX
CC   The present invention relates to a recombinant phytase (EC 3.1.2.8)
CC   protein and its encoding also nucleic acid molecule. Phytases of the
CC   invention is a thermostable protein. They are useful in foodstuffs, for
CC   oil degumming, producing an animal feed, delivering a phytase enzyme
CC   supplement to an animal, increasing the resistance of a phytase
CC   polypeptide to enzymatic inactivation in a digestive system of an animal
CC   and in processing of corn and sorghum kernels. The invention is also

```
CC   useful in gene therapy and in production of transgenic animals. The
CC   present sequence is a Escherichia coli phytase appA gene. Note: The
CC   description specified in the specification for the SEQ ID NO:3 is E. coli
CC   phytase appA gene. However the description specified in the accession
CC   number M58708 is E. coli periplasmic phosphoanhydride phosphohydrolase
CC   (appA) gene. Thus the description given in the specification varies with
CC   the accession number.
XX
SQ   Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;

Query Match      97.4%;  Score 1287.8;  DB 14;  Length 1901;
Best Local Similarity   99.5%;  Pred. No. 0;
Matches 1293;  Conservative   0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy        1 ATGAAAACGGATCTTAATCCCATTTTTATCTCTTCTGATTCCTTAACCCCGCAAATCTGCA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      188 ATGAAAACGGATCTTAATCCCATTTTTATCTCTTCTGATTCCTTAACCCCGCAAATCTGCA 247

Qy       61 TTCGCTCAGAGTGAGCGCGGAGCTGAAGCTGAAAAGTGTGGTGATTGTCAGTGTCATGGT 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      248 TTCGCTCAGAGTGAGCGCGGAGCTGAAGCTGAAAAGTGTGGTGATTGTCAGTGTCATGGT 307

Qy      121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCCATGGCCA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCCATGGCCA 367

Qy      181 ACCTGGCCGGGTAAAACTGGGTTGGGTCGGCTGAGCTGACCATAACCCGACTC 240
            |||||||||||||||||||||||||||||||||||||||||  |||||||||||
Db      368 ACCTGGCCGGGTAAAACTGGGTTGGGTCGGCTGAGCTAATCGGCCTATCTC 427

Qy      241 GGACATTACCAACGCCAGCGTGTCGGTAGCCGACGGAGAAAAGGCTGCCCG 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      428 GGACATTACCAACGCCAGCGTCGGTAGCCGACGGAGAAAAGGCTGCCCG 487

Qy      301 CAGTCTGGTCAGGTCGCGGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      488 CAGTCTGGTCAGGTCGCGGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA 547

Qy      361 GCCTTCGCGCGCCGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      548 GCCTTCGCGCGCCGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 607

Qy      421 TCCCAGTCCCGGATCCGGTTATTTAATCCTCTAAAAAACTGGCGTTTGCCAACTGATAACGCG 480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      608 TCCCAGTCCCGGATCCGGTTATTTAATCCTCTAAAAAACTGGCGTTTGCCAACTGATAACGCG 667

Qy      481 AACGTGACTGACGCGATCCTCAGCAGGGCAGAGAGGGTCAAGTGCTGACTTTACCGGGCAT 540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      668 AACGTGACTGACGCGATCCTCAGCAGGGCAGAGAGGGTCAAATGCTGACTTTACCGGGCAT 727

Qy      541 CCGGCAAAACGGCCGTTTCGCGAACTGGGACGGTCGTTAATTTTCCGCAAATCAAACTTGTTGC 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      728 CCGGCAAACGGCGTTTCGCGGAACTGGGACGGTCGTTAATTTTCCGCAAATCAAACTTGTTGC 787

Qy      601 CTTAAAACGTGAGAAAACAGGACGAAAAGCTGTTCATTAACGCAGGACATTACCATCGGAACTC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      788 CTTAAACGTGAGAAAACAGGACGAAAAGCTGTTCATTAACGCAGGACATTACCATCGGGAACTC 847

Qy      661 AAGGTGAGCGCCGACACAATGTCTTCATTAAACCGGTGCGGTAAGCCTTGCGATCAATGCTGACG 720
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      848 AAGGTGAGCGCCGACAATGTCTTCATTAACCGGTGCGGTAAGCCTTGCGATCAATGCTGACG 907

Qy      721 GAGATATTTCTCCTGCAACACAGCGGATTTTTCAACACAGCAGGGAATCC 780
            |||||||||||||||||||| || ||||||||||||||||||||||||||
Db      908 GAGATATTTCTCCTGCAACACAGCGGAATGCCGGAGCCGGGTGGGGAGGGATCACC 967

Qy      781 GATTCACCAGTGGAACACCTTGCTAAGTTGCATAACGCGCGCAATTTTATTTTGCTACAA 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db      968 GATTCACACCAGTGGAACACCTTGCTAAGTTGCATAACGCGCGCAATTTTATTTTGCTACGAG 1027

Qy      841 CGCACGCCAGAGGTTGCCCGCAGCCGCGGCCACCCCGGTTATTAGATTTGATCAAGGCAGCG 900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1028 CGCACGCCAGAGGTTGCCCGCAGCCGCGGCCACCCCGGTTATTAGATTTGATCAAGACAGCG 1087

Qy      901 TTGACGCGCCCATCCACCGCAAAAACAGGCGTATGGGTCGTGACATTACCCACTTCAGTACTG 960
            ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||| |||
Db     1088 TTGACGCGCCCATCCACCGCAAAAACAGGCGTANGGGTGTGACATTACCCACTTCAGTGCTG 1147

Qy      961 TTTATTGCCGGACACGATACTRAATCTGGCAAAATCTCGGCGGCCCAGGCCACTGGAGTCAACTGG 1020
            ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1148 TTTATCGCCGGACACGATACTRAATCTGGCAAAATCTCGGCGGCCCAGGCCACTGGAGTCAACTGG 1207

Qy     1021 ACGGCTTCCCGGTCAGCCGCGATAAACACCCCGCCCAGGGCTGGTGAACTGTTGGTTTGAAACGCTGG 1080
            |||| ||||||||||||||||  |||||||  ||||||||||||||||||||||| |||||||||||||
Db     1208 ACGCTTCCCGGTCAGCCGGATAAACACGGCGCCCAGGGCTGGTGAACTGGTGTTTGAAACGGCTGG 1267

Qy     1081 CCTCGGCTTAAGGCCGATAACACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1268 CGTCGGCTTAAGGCCGATAACACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1327

Qy     1141 CAGATGCCTGATAAAACGCCGCTGTCATTAAATATACGGCGCCCGGAGAGGTGAAACTGACC 1200
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1328 CAGATGCCTGATAAAACGCCGCTGTCATTAAATATACGGCGCCCGGAGAGGTGAAACTGACC 1387

Qy     1201 CTGGGCACGATGTGAAGAAGAGCGAAATGCCGCCAGGGCCGCCCATGTTCGTTGTTCGTTGGCAGGGTTTTACGGCAA 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1388 CTCGGCACGATGTGAAGAAGACCGAAATGCCGCCAGGGCCATGTTCGTTGGGCAGGGTTTTTACGGCA 1447

Qy     1261 ATCGTGAATGAAGCACGGCCATACCCGGCGGTGCAGTTTGAGAT 1300
            ||||||||||||||||||||||||||||||||||||||||  ||
Db     1448 ATCGTGAATGAAGCACCGCCATACCGGGCGGTGCAGTTTGTGTAAT 1487

RESULT 14
AEE75418

ID   AEE75418 standard; DNA; 1901 BP.
XX
AC   AEE75418;
XX
DT   23-FEB-2006  (first entry)
XX
DE   E. coli B phytase appA coding sequence SEQ ID NO: 7.
XX
KW   ds; gene; thermostable; phytase; protein stabilization; pharmaceutical;
KW   fertilizer; cramp; cell culture; osteoporosis; osteopathic; food; appA.
OS   Escherichia coli K12.
XX
FH   Key            Location/Qualifiers
FH
FT   CDS            188..1483
FT                  /*tag= a
FT                  /product= "Phytase"
FT   unsure         401..403
FT                  /*tag= b
FT                  /note= "ecodes Arg"
XX
PN   US2005281792-A1.
XX
PD   22-DEC-2005.
XX
PF   01-SEP-2004; 2004US-00933115.
XX
PR   13-AUG-1997; 97US-00910798.
PR   01-MAR-1999; 99US-00259214.
PR   13-APR-1999; 99US-00291931.
PR   25-MAY-1999; 99US-00318528.
PR   25-MAY-2000; 2000US-00580515.
PR   24-MAY-2001; 2001US-00866379.
XX
PA   (SHOR/) SHORT J M.
PA   (KRET/) KRETZ K A.
PA   (GRAY/) GRAY K A.
PA   (BART/) BARTON N R.
PA   (GARR/) GARRETT J B.
PA   (ODON/) ODONOGHUE E.
PA   (BAUM/) BAUM W.
```

```
PA   (ROBE/) ROBERTSON D E.
PA   (ZORN/) ZORNER P.
XX
PI   Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, Odonoghue E;
PI   Baum W, Robertson DE, Zorner P;
XX
DR   WPI; 2006-055022/06.
DR   P-PSDB; AEF75419.
XX
PT
PT
PT
XX
PS   Claim 1; SEQ ID NO 7; 82pp; English.
XX
CC   The present sequence is that of a polynucleotide encoding the Escherichia
CC   coli phytase appA, which was modified to produce a thermostable protease
CC   resistant phytase which constitutes the novel enzyme of the invention.
CC   The present invention relates to a novel formulation, comprising at least
CC   one polypeptide having phytase activity, where the polypeptide is a
CC   modified sequence (AEF75421) derived from the Escherichia coli K-12 appA
CC   phytase (AEF75419). The modification of the enzyme, comprising a series
CC   of mutations, improves the thermal tolerance and protease stability of
CC   the protein. The specification also claims a pharmaceutical composition;
CC   a kit; an immobilized phytase; a fertilizer or soil additive; a liquid
CC   supplement for preventing muscle cramps; a hydrating agent; a tissue
CC   culture or cell culture media; and a plant food additive, all comprising
CC   the novel phytase, and a method of reducing pollution and increasing
CC   nutrient availability in an environment or environmental sample by
CC   degrading environmental phytic acid. The novel enzyme is a dietary
CC   supplement useful for treating, preventing or reversing osteoporosis or
CC   bone loss, and preventing muscle cramps. The liquid supplement is useful
CC   for preventing muscle cramps. The method is useful for reducing pollution
CC   and increasing nutrient availability in an environment or environmental
CC   sample by degrading environmental phytic acid, where the environment or
CC   environmental sample comprises a soil or a body of water. The immobilized
CC   phytase is useful in foodstuffs for improving the feeding value of
CC   phytate rich ingredients.
XX
SQ   Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;

Query Match             99.4%;  Score 1287.8;  DB 15;  Length 1901;
Best Local Similarity   97.4%;
Matches 1293;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      1 ATGAAAGCGGATCTTAATCCATTTTATCTCTTCTTCCGTTAACCCGCAATCTGCA     60
          |||||| ||||||||||||| |||||||||||||||  ||||||||||||||||||
Db    188 ATGAAACGGATCTTAATCCCATTTTATCTCTTCTGATTCCGTTAACCCGCAATCTGCA   247

Qy     61 TTCGCTCAGAGTGAGCCGGAGTTGAAGTGGAGAAAGTGTGGTCAGTCGTCTCATGGT    120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248 TTCGCTCAGAGTGAGCCGGAGTTGAAGTGGAGAAAGTGTGGTCAGTCGTCTCATGGT    307

Qy    121 GTGGCGTTGCTCAACCAAGGCCACGCAACTGATCGCAGGATGTCACCCCAGACCGCA    180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308 GTGCGTGCCTCCAACCAAGGCCACGCAACTGATCGCAGGATGTCACCCCAGACGCATGGCCA  367

Qy    181 ACCTGGCCGGGTAAAACTGGGTTGGCGTGAGCGCTAATCGCCCTATCTC    240
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    368 ACCTGGCCGGGTAAAACTGGGTTGGCGTGAGCGCTAATCGCCCTATCTC    427

Qy    241 GGACATTACCAACGCCAGCGTCTGGCTAGCCGACGGATTGCTCGGCGAAAAAGGGCTGCCCG   300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428 GGACATTACCAACGCCAGCGTCTGGCTAGCCGACGGATTGCTCGGCGAAAAAGGGCTGCCCG   487

Qy    301 CAGTCTGGTCAGGTGCGCCGATTATTGCTGATGTCGACGAGGCGTACCCGTAAAACAGGCCAA   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488 CAGTCTGGTCAGGTGCGCCGATTATTGCTGATGTCGACGAGGCGTACCCGTAAAACAGGCCAA   547

Qy    361 GCCTTCGCGCCGCGGGCTGGCACCTGACTGTGCAATAACCGTACATAACCCAGGCCAGATACG    420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548 GCCTTCGCGCCGCGGGCTGGCACCTGACTGTGCAATAACCGTACATAACCCAGGCCAGATACG    607

Qy    421 TCCAGTCCCGGATCCGTTATTAATCCTCTAAAAACTGGGCGGTTGCCAACTGGGATAACGCG    480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608 TCCAGTCCCGGATCCGTTATTAATCCTCTAAAAACTGGGCGGTTGCCAACTTACCGGCAT    667

Qy    481 AACCTGACTGACGGCGATCCTCCAGCAGGGCGAGGGAGGGGTCAATTGCTGACTTTACCGGCAT    540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    668 AACCTGACTGACGGCGATCCTCCAGCAGGGCGAGGGAGGGGTCAATTGCTGACTTTACCGGCAT    727

Qy    541 CGGCAAAACGGCGTTTCGGCGAACTGGAAACGGGTGCTTAATTTTCCCGCAATCAAACTTGTGC    600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    728 CGGCAAAACGGCGTTTCGGCGAACTGGAAACGGGTGCTTAATTTTCCCGCAATCAAACTTGTGC    787

Qy    601 CTTAAACGTGAGAAAACAGGACGGAGAAAGCTGTTCATTAACGGCAGGCAGGAACTC    660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    788 CTTAAACGTGAGAAAACAGGACGGAGAAAGCTGTTCATTAACGGCAGGCAGGAACTC    847

Qy    661 AAGGTGAGGCGCCGACAAATGTCTCATTAACGGCGTCGGTAAGCCTCGGCATTCAAATGCTGACG   720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    848 AAGGTGAGGCGCCGACAAATGTCTCATTAACGGCGTCGGTAAGCCTCGGCATTCAAATGCTGACG   907

Qy    721 GAGATATTTCTCCTGCAACAAGCACAGGGAAATGCGGAGCGGGGGGTGGGGAAGGATCACC    780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908 GAGATATTTCTCCTGCAACAAGCACAGGGAAATGCGGAGCGGGGGGTGGGGAAGGATCACC    967

Qy    781 GATTCACCACCAGTGGAAACACCTTGCTAAGTTGCAATAAAGGGCAATTTATTTATTTGCTACAA   840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968 GATTCACCACCAGTGGAAACACCTTGCTAAGTTGCAATAAAGGGCAATTTATTTATTTGCTACAA   1027

Qy    841 CGCACGCCAGAGGTTGCCGCCAGCCCGCGCGCACCCCGTTAATTTGGATTTGATCATGGCACGCG   900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028 CGCACGCCAGAGGTTGCCGCCAGCCCGCGCGCACCCCGTTAATTTGGATTTGATCAAGAGACAGCG   1087

Qy    901 TTGACGCCCCATCCACCGCAAAAACAGGGGTAAGGTAAGTGAGACATTACCCACTTCAGTACTG    960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 TTGACGCCCCATCACCGCAAAAACAGGGGTATGGTGTGACATTACCCACTTCAGTGCTG   1147

Qy    961 TTTATTGCCGGACACGGATACTAATCTGGGAAAATCTCGGGCGGCGGCACTGGAGCTCAAACTGG   1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 TTTATCGCGCGGACACGGATACTAATCTGGGAAAATCTCGGGCGGCGGCACTGGAGCTCAAACTGG   1147

Qy   1021 ACGCTTCCCGGTCAGCCCGGATAACACGCGCCGGCGCCACTGGTGGTAAACTGGTGTTTGAACGGCTGG   1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148 ACGCTTCCCGGTCAGCCCGGATAACACGCGCCGGCGCCACTGGTGGTAAACTGGTGTTTGAACGGCTGG   1207

Qy   1081 CGTCGGGCTAAGCGGATAACAGCCAGTGGATTCAGGTTTCGGCTGTCTTCCAGACTTTACAG   1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208 CGTCGGGCTAAGCGGATAACAGCCAGTGGATTCAGGTTTCGGCTGTCTTCCAGACTTTACAG   1267

Qy   1141 CAGATGCGTGATAAAAGCGCTGTCATTAAAATAGCGCGCCCCGGAGAGGTGAAAACTGACC   1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268 CAGATGCGTGATAAAAGCGCTGTCATTAAAATAGCGCGCCCCGGAGAGGTGAAAACTGACC   1327

Qy   1201 CTGGCAGGATGTGAAGAGCGGAAATGCCAGGGCCATGTGTTCGTTGGCAGGTTTACGCAA   1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328 CTGGCAGGATGTGAAGAGCGGAAATGCCAGGGCCATGTGTTCGTTGGCAGGTTTACGCAA   1387

Qy   1261 ATCCGTGAATGAAGGACGACGCATACCCGCGTGCAGTTTGTGAGAT   1300
          ||||||||||||||||||||||||||||||||||||||||||| |
Db   1388 CTGGCAGGATGTGAAGAGCGGAAATGCCAGGGCCATGTGTTCGTTGGCAGGTTTACGCAA   1447

Db   1448 ATCCGTGAATGAAGAGCACGCATACCCGCGTGCAGTTTGTGAAT   1487
```

```
RESULT 15
ABK12514
ID   ABK12514 standard; DNA; 1299 BP.
XX
AC   ABK12514;
XX
DT   05-JUN-2002  (first entry)
XX
DE   DNA encoding phytase associated protein.
XX
KW   Phytase; gene; ds.
XX
```

```
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1299
FT                   /*tag= a
FT                   /product= "Phytase associated protein"
XX
PN   KR9086028-A.
XX
PD   15-DEC-1999.
XX
PF   25-MAY-1998;  98KR-00018810.
XX
PR   25-MAY-1998;  98KR-00018810.
XX
PA   (WOOJ-) WOOJIN CO LTD.
XX
PI   Bae HD,  Forceburgh CW,  Goloben S,  Cheng KJ;
XX
DR   WPI; 2000-645078/62.
DR   P-PSDB; AAU77775.
XX
PT   Novel phytase gene, recombinant phytase and usage thereof.
XX
PS   Claim 1; Fig 2; 10pp; Korean.
XX
CC   The invention relates to a novel phytase gene, a recombinant phytase gene
CC   and their uses. This sequence encodes a phytase associated protein,
CC   described in the invention
XX
SQ   Sequence 1299 BP; 317 A; 345 C; 355 G; 282 T; 0 U; 0 Other;

Query Match           97.4%;  Score 1287;  DB 3;  Length 1299;
Best Local Similarity 99.5%;  Pred. No. 0;
Matches 1290;  Conservative   0;  Mismatches   6;  Indels   0;  Gaps   0;
```

```
Qy       1 ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTCTCTGATTCGTTAACCCCGCAAATCTGCA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTCTCTGATTCGTTAACCCCGCAAATCTGCA 60

Qy      61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGAAAGTGTGGTGATTGTCAGTCGTGTCATCAGGT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGAAAGTGTGGTGATTGTCAGTCGTGTCATCAGGT 120

Qy     121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACCGCCATGGCCA 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACCGCCATGGCCA 180

Qy     181 ACCTGGCCGGTAAAACTGGGTTGGCTGACGACACCGGCGGTGTGGAGCTAATCGCCTATCTC 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ACCTGGCCGGTAAAACTGGGTTGGCTGACGACACCGGCGGTGTGGAGCTAATCGCCTATCTC 240

Qy     241 GGACAATTACCAAACGCCAGCGTCTGGGTAGCCGACGGATTGCTGGGCGAAAAAGGGCTGCCCG 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GGACAATTACCAAACGCCAGCGTCTGGGTAGCCGACGGATTGCTGGGCGAAAAAGGGCTGCCCG 300

Qy     301 CAGTCTGGTCAGTCGCGGATTATTGCTGAGTGCAGGAGCGTACCCGTAAAACAGGCGAA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CAGTCTGGTCAGTCGCGGATTATTGCTGAGTGCAGGAGCGTACCCGTAAAACAGGCGAA 360

Qy     361 GCCTTCGCCGCCGCGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGCGCAGATACG 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GCCTTCGCCGCCGCGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGCGCAGATACG 420

Qy     421 TCCAGTCCGATCCGTTATTTAATCCTCTAAAAAACTGGCGTTTGCCAACTGGATAACGCGG 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TCCAGTCCGATCCGTTATTTAATCCTCTAAAAAACTGGCGTTTGCCAACTGGATAACGCGG 480

Qy     481 AACGTGACTGACGCGGATCCTCAGCAGGGCAGGAGGGGTCAAATTGCTGACTTTACCGGGCAT 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AACGTGACTGACGCGGATCCTCAGCAGGGCAGGAGGGGTCAAATTGCTGACTTTACCGGGCAT 540

Qy     541 CGGCAAACGGCGGTTTCCGACTGGAACGGGGTGCTTAATTTTTCCGCAATCAAACTTGTGC 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 CGGCAAACGGCGGTTTCCGACTGGAACGGGGTGCTTAATTTTTCCGCAATCAAACTTGTGC 600

Qy     601 CTTAAACGGTGAGAAAACAGGAGGAGAAAAGCTGTTCATTAACGCAGGCACATTACCATCG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 CTTAAACGGTGAGAAAACAGGAGGAGAAAAGCTGTTCATTAACGCAGGCACATTACCATCG 660

Qy     661 AAGGTGAGCGCCGGACAAATGTTCTCATTAACCGGTGCGGTAAGCCTTCGGCATCGAATGCTCGACG 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AAGGTGAGCGCCGGACAAATGTTCTCATTAACCGGTGCGGTAAGCCTTCGGCATCGAATGCTCGACG 720

Qy     721 GAGATATTTCTCCTGCAACAACGACCAGGAGAAGGATACCCGGGCCGGGCGTGGGGAGGATCACC 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GAGATATTTCTCCTGCAACAACGACCAGGAGAAGGATACCCGGGCCGGGCGTGGGGAGGATCACC 780

Qy     781 GATTCACACCAGTGGAACACTTGCTAAGTTTGCATAAACGCCAATTTTATTTGCTACAA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GATTCACACCAGTGGAACACTTGCTAAGTTTGCATAAACGCCAATTTTATTTGCTACAA 840

Qy     841 CGCACGCCAGAGGTTGCCCGGCCAGCCGGCCCACCCCGTTATTGGATTTGATCATGGCAGCG 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CGCACGCCAGAGGTTGCCCGGCCAGCCGGCCCACCCCGTTATTGGATTTGATCATGGCAGCG 900

Qy     901 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGGTGTGACATTACCCACTTCAGTGCTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGGTGTGACATTACCCACTTCAGTGCTG 960

Qy     961 TTTATTGCCCGGACACGATACTAATTCTCGGCAAAATCTCGGCGGGGCCACTGGAGCTCAACTGG 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 TTTATTGCCCGGACACGATACTAATTCTCGGCAAAATCTCGGCGGGGCCACTGGAGCTCAACTGG 1020

Qy    1021 ACGCTTCCCGGTTCAGCGCGGATAAACACCCCGCCAGGTGGTGAAACTGGTTGTTGTTTGAAACGCTGG 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 ACGCTTCCCGGTTCAGCGCGGATAAACACCCCGCCAGGTGGTGAAACTGGTTGTTGTTTGAAACGCTGG 1080

Qy    1081 CGTCGGCGTAAGCCGATAACAGCCAGTGGGATTCAGGTTTCGGCTGGTTCTTCCAGACTTTACAG 1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CGTCGGCGTAAGCCGATAACAGCCAGTGGGATTCAGGTTTCGGCTGGTTCTTCCAGACTTTACAG 1140

Qy    1141 CAGATGGCTGATAAAACCCGCTGTCATTAAATACCGCCGCCCCGGAGAGGTGAAACTGACC 1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 CAGATGGCTGATAAAACCCGCTGTCATTAAATACCGCCGCCCCGGAGAGGTGAAACTGACC 1200

Qy    1201 CTTGGCAGGATGTTGAAGAGGCGGAAATGCCCAGGCGCCATGTTGTTCGTTGGCAGGTTTTACCGCAA 1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 CTTGGCAGGATGTTGAAGAGGCGGAAATGCCCAGGCGCCATGTTGTTCGTTGGCAGGTTTTACCGCAA 1260

Qy    1261 ATCGTGAATGAATGAAAGCACGGCCATACCGGCGTGCCAGTTTG 1296
           ||||||||||||||||||||||||||||||||||||||||||||
Db    1261 ATCGTGAATGAATGAAAGCACGGCCATACCGGCGTGCCAGTTTG 1296
```

```
Search completed: June 14, 2006, 10:31:32
Job time : 903.103 secs
```

OM nucleic - nucleic search, using sw model

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

Run on: June 14, 2006, 10:00:55 ; Search time 7399.95 Seconds
               (without alignments)
               11432.839 Million cell updates/sec

Title: US-10-601-319-1
Perfect Score: 1322
Sequence: 1 atgaaagcgatcttaatccc..........atcaccatcaccactaa 1323

Scoring table: IDENTITY NUC
                Gapop 10.0 , Gapext 1.0
Searched: 6366136 seqs, 3197710525 residues
Total number of hits satisfying chosen parameters: 0
Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
                 Maximum Match 100%   12732272
                 Listing first 45 summaries

Database :   GenEmbl:*
                1: gb_env:*
                2: gb_pat:*
                3: gb_ph:*
                4: gb_pl:*
                5: gb_pr:*
                6: gb_ro:*
                7: gb_sts:*
                8: gb_sy:*
                9: gb_un:*
                10: gb_vi:*
                11: gb_ov:*
                12: gb_htg:*
                13: gb_in:*
                14: gb_om:*
                15: gb_ba:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1322	100.0	1323	2	AR108133	AR108133 Sequence
2	1322	100.0	1323	2	AR127818	AR127818 Sequence
3	1322	100.0	1323	2	AR130956	AR130956 Sequence
4	1322	100.0	1323	2	AR492966	AR492966 Sequence
5	1322	100.0	1323	2	AR636185	AR636185 Sequence
6	1322	100.0	1323	2	AX052713	AX052713 Sequence
7	1322	100.0	1323	2	AX338966	AX338966 Sequence
8	1322	100.0	1323	2	AX356566	AX356566 Sequence
9	1287.8	97.4	1901	2	AR636190	AR636190 Sequence
10	1287.8	97.4	1901	2	AX356572	AX356572 Sequence
11	1287.8	97.4	1901	2	ECOAPPAA	M58708 Escherichia
12	1287.8	97.4	110000	15	U00096_10	Continuation (11 o
13	1287.8	97.4	1901	2	AP009048_10	Continuation (11 o
14	1286.2	97.3	1901	2	AR636188	AR636188 Sequence
15	1286.2	97.3	1901	2	AR636191	AR636191 Sequence
16	1286.2	97.3	1901	2	AX356570	AX356570 Sequence
17	1286.2	97.3	1901	2	AX356574	AX356574 Sequence
18	1286.2	97.3	1901	15	ECOACPHB	L03371 Escherichia
19	1286.2	97.3	110000	15	CP000038_10	Continuation (11 o
20	1284.6	97.2	1315	15	AF537219	Escherich
21	1284.6	97.2	1901	15	L03372	Escherichia
22	1284.6	97.2	1901	15	L03373	Escherichia
23	1284.6	97.2	1901	15	L03374	Escherichia
24	1284.6	97.2	1901	15	L03375	Escherichia
25	1283	97.0	1901	15	ECOACPHF	L03375 Escherichia
26	1283	97.0	1901	15	ECOACPHE	AR636189 Sequence
27	1283	97.0	1901	15	ECOACPHA	L03370 Escherichia
28	1283	97.0	110000	15	CP000036_22	Continuation (23 o
29	1282.2	97.0	3470	2	AX042376	AX042376 Sequence
30	1282.2	97.0	4060	2	AX042374	AX042374 Sequence
31	1282.2	97.0	6116	2	AX042375	AX042375 Sequence
32	1282.2	97.0	6708	2	AX042373	AX042373 Sequence
33	1282.2	97.0	17732	2	AX042378	AX042378 Sequence
34	1282.2	97.0	20623	2	AX042372	AX042372 Sequence
35	1273.4	96.3	110000	15	AE005674_10	Continuation (11 o
36	1271.8	96.2	110000	15	AE014073_10	Continuation (11 o
37	1270.2	96.1	5421	2	AX277753	AX277753 Sequence
38	1263.8	95.6	1489	2	AX042377	AX042377 Sequence
39	1263.8	95.6	6708	2	AY496073	AY496073 Escherich
40	1263	95.5	1486	2	AR630587	AR630587 Sequence
41	1251.8	94.7	1486	2	AR630588	AR630588 Sequence
42	1246.6	94.3	110000	15	AE005174_13	Continuation (13 o
43	1246.6	94.3	110000	15	AE005174_12	Continuation (13 o
44	1246.6	94.3	110000	15	BA000007_12	Continuation (13 o
45	1246	94.3	110000	15	AB040075_10	Continuation (11 o

ALIGNMENTS

RESULT 1
AR108133

LOCUS   AR108133   1323 bp   DNA   linear   PAT 14-FEB-2001
DEFINITION Sequence 1 from patent US 6110719.
ACCESSION   AR108133
VERSION   AR108133.1   GI:12823620
KEYWORDS   .
SOURCE   Unknown.
ORGANISM   Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1323)
AUTHORS   Kretz, K.
TITLE   Phytase
JOURNAL   Patent: US 6110719-A 1 29-AUG-2000;
FEATURES   Location/Qualifiers
source   1..1323
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Query Match   100.0%; Score 1322; DB 2; Length 1323;
Best Local Similarity   100.0%;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCGTTAACCCCGCAATCTGCA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCGTTAACCCCGCAATCTGCA 60

Qy   61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120

Qy  121 GTGGCGTGCTCCAACCAAGCCACGCAACTGATGATGCAGGATGCATGGCCA 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTGGCGTGCTTCCAACCAAGCCACGCAACTGATGATGCAGGACATGGCCA 180

Qy  181 ACCTGGCGCCGGTAAAACTGGTTGGCTGACACCCGCGNGGGTGGTGAGCTAATGCGCTATCTC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ACCTGGCGCCGGTAAAACTGGTTGGCTGACACCCGCGGGTGGTGAGCTAATCGCCTATCTC 240

Qy  241 GGACATTACCAACGCCAGCGTCGTGGTAGGCACGACGGATTGCTGCGGAAAAAGGGGCTGCCCG 300
```

```

```
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG 300

Qy   301 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA 360

Qy   361 GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420

Qy   421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG 480

Qy   481 AACGTGACTGACGCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AACGTGACTGACGCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 540

Qy   541 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600

Qy   601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC 660

Qy   661 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 720

Qy   721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGTGGGGAAGGATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGTGGGGAAGGATCACC 780

Qy   781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA 840

Qy   841 CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTGGATTTGATCATGGCAGCG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTGGATTTGATCATGGCAGCG 900

Qy   901 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTACTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTACTG 960

Qy   961 TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG 1020

Qy   1021 ACGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1080

Qy   1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140

Qy   1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCCGGAGAGGTGAAACTGACC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCCGGAGAGGTGAAACTGACC 1200

Qy   1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA 1260

Qy   1261 ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320

Qy   1321 TAA 1323
              |||
Db   1321 TAA 1323

RESULT 2
AR127818
LOCUS   AR127818   1323 bp   DNA   linear   PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6183740.
ACCESSION   AR127818
VERSION   AR127818.1 GI:14115480
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1323)
  AUTHORS   Short,J.M. and Kretz,K.A.
  TITLE   Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: US 6183740-A 1 06-FEB-2001;
FEATURES   Location/Qualifiers
     source   1. .1323
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match   100.0%; Score 1322; DB 2; Length 1323;
  Best Local Similarity   100.0%; Pred. No. 0;
  Matches 1323; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60

Qy   61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120

Qy   121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy   181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC 240

Qy   241 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG 300

Qy   301 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA 360

Qy   361 GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420

Qy   421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG 480

Qy   481 AACGTGACTGACGCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AACGTGACTGACGCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 540

Qy   541 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600

Qy   601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC 660

Qy   661 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Db 661 AAGGTGAGCGGCCGACAAATGTCTCATTAACCGGTGCGGTGGGTAAGCCTCGCATCAAGTGCTGACG 720

Qy 721 GAGATATTTCTCCTGCAACAAGCACAGGGAAATGCCGGACGCCGGGGGTGGGGAAGGATCACC 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 721 GAGATATTTCTCCTGCAACAAGCACAGGGAAATGCCGGAGCCGGGGGTGGGGAAGGATCACC 780

Qy 781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCGAATTTTATTTGCTACAA 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCGAATTTTATTTGCTACAA 840

Qy 841 CGCACGCCAGAGGTTGCCCGCGCCGCGCCACCCCGTTATTGGATTTGATCATGGCAGCG 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 841 CGCACGCCAGAGGTTGCCCGCGCCGCGCCACCCCGTTATTGGATTTGATCATGGCAGCG 900

Qy 901 TTGACGCCCCATCCACCCGCAAAAACAGGGGTATGGTGTGACATTACCCACTTCAGTACTG 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 901 TTGACGCCCCATCCACCCGCAAAAACAGGGGTATGGTGTGACATTACCCACTTCAGTACTG 960

Qy 961 TTTATTGCCGGACACGATACTAATCTCGGCAAAATCTCGGCGGAGCTCAAACTGG 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 961 TTTATTGCCGGACACGATACTAATCTCGGCAAAATCTCGGCGGAGCTCAAACTGG 1020

Qy 1021 ACGCTTCCCGGTCAGCCGGATAACACGCCGCCCAGGTGGGTGAACTGGTTTGAACGCTGG 1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 ACGCTTCCCGGTCAGCCGGATAACACGCCGCCCAGGTGGGTGAACTGGTTTGAACGCTGG 1080

Qy 1081 CGTCGGGCTAAGCCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCAGACTTTACAG 1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 CGTCGGGCTAAGCCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCAGACTTTACAG 1140

Qy 1141 CAGATGCCTGATAAAAACCGCGCGTGTCATTAATACGCGCCCCGGAGGAGGGTGAAAACTGACC 1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 CAGATGCCTGATAAAAACCGCGCGTGTCATTAATACGCGCCCCGGAGGAGGGTGAAAACTGACC 1200

Qy 1201 CTGGCAGGATGTGAAGAGAGGCGAAAATGCGCAGGGGCCATGCTGGTTTGTTCGTTCGTTGGCAGGTTTTACGGCAA 1260
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 CTGGCAGGATGTGAAGAGAGGCGAAAATGCGCAGGGGCCATGCTGGTTTGTTCGTTCGTTGGCAGGTTTTACGGCAA 1260

Qy 1261 ATCGTGAATGAAGCGACGCATACCCGGCGGCTGCCAGTTTGAGATCTCATCACCATCACCATCAC 1320
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 ATCGTGAATGAAGCGACGCATACCCGGCGGCTGCCAGTTTGAGATCTCATCACCATCACCATCAC 1320

Qy 1321 TAA 1323
        |||
Db 1321 TAA 1323
```

RESULT 3

```
LOCUS       AR130956                1323 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 1 from patent US 6190897.
ACCESSION   AR130956
VERSION     AR130956.1  GI:14119281
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1323)
  AUTHORS   Kretz,K.
  TITLE     Phytase
  JOURNAL   Patent: US 6190897-A 1 20-FEB-2001;
FEATURES             Location/Qualifiers
     source          1..1323
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN
```

```
Query Match             100.0%;  Score 1322;  DB 2;  Length 1323;
Best Local Similarity   100.0%;
Matches 1323;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy    1 ATGAAAAGGGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCGCGAATCTGCA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAAGGGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCGCGAATCTGCA 60
```

```
Qy   61 TTCCGTCTCAGAGTGAGCCGGAGCTCGAAGCTGAAAGCTTGGAAAGTGATTGTCAGTCGGT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GTGCGTGCTTCCAACCAAGGCCACGCAACTGATGCAGAAGCTCAGTCGGT 120

Qy  121 GTGCGTGCTTCCAACCAAGGCCACGCAACTGATGCAGAAGCTCAGTCGGT 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GGACATTACCAACCGCCAGTGGTTCTCGGTTAGCCGACGGATTGCTGGGGAAAAAGGGCTGCCCG 180

Qy  181 ACCTGGCCCGGTAAAAACTGGGTTGGCGTGACCGGCACCGCGGTAAGCTAACGCGAA 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ACCTGGCCCGGTAAAAACTGGGTTGGCGTGACCGGCACCGCGGTAAGCTAACGCGAA 240

Qy  241 GGACATTACCAACCGCCAGTGGTTCTCGGTTAGCCGACGGATTGCTGGGGAAAAAGGGCTGCCCG 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GGACATTACCAACCGCCAGTGGTTCTCGGTTAGCCGACGGATTGCTGGGGAAAAAGGGCTGCCCG 300

Qy  301 CAGTCTGGTCAGGTCGGCGATTAATTGCTGATGTCGACGGAGGGTACCGGTAAAACAGGCGAA 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGTCTGGTCAGGTCGGCGATTAATTGCTGATGTCGACGGAGGGTACCGGTAAAACAGGCGAA 360

Qy  361 GCCTTCGCGCGGCCGGGCTGTGCCAACCTGACTGTGCAATAAACCCTACATACCAGGCAGATACG 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GCCTTCGCGCGGCCGGGCTGTGCCAACCTGACTGTGCAATAAACCCTACATACCAGGCAGATACG 420

Qy  421 TCCAGTCCCGATCCGTCGTTATTTAATCCTCTAAAAAACTGGCGGTTTGCCGGATAACGCG 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TCCAGTCCCGATCCGTCGTTATTTAATCCTCTAAAAAACTGGCGGTTTGCCGGATAACGCG 480

Qy  481 AACGTGACTGAACGGATCCTTCAGCAGAGGGAGGAGGGTCAAATGCTGCTGACTTTACCGGGCAT 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 AACGTGACTGAACGGATCCTTCAGCAGAGGGAGGAGGGTCAAATGCTGCTGACTTTACCGGGCAT 540

Qy  541 CGGCAAAACCGGCGTTTCGCGGAAACTGGAACCTGAAACGGTCGTTAATTTCCCGCAATCAAACTTGTGC 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 CGGCAAAACCGGCGTTTCGCGGAAACTGGAACCTGAAACGGTCGTTAATTTCCCGCAATCAAACTTGTGC 600

Qy  601 CTTAAACGTGAGAAAACAGGACGAAAAGCTGTTCATTAACGCAGGCATTACCATTACCATCCGGAACTC 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CTTAAACGTGAGAAAACAGGACGAAAAGCTGTTCATTAACGCAGGCATTACCATTACCATCCGGAACTC 660

Qy  661 AAGGTGAGCGGCCGACAAATGTCTCATTAACCGGTGCGGTGGGTAAGCCTCGCATCAAGTGCTGACG 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 AAGGTGAGCGGCCGACAAATGTCTCATTAACCGGTGCGGTGGGTAAGCCTCGCATCAAGTGCTGACG 720

Qy  721 GAGATATTTCTCCTGCAACAAGCACAGGGAAATGCCGGACGCCGGGGGTGGGGAAGGATCACC 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GAGATATTTCTCCTGCAACAAGCACAGGGAAATGCCGGAGCCGGGGGTGGGGAAGGATCACC 780

Qy  781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCGAATTTTATTTGCTACAA 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCGAATTTTATTTGCTACAA 840

Qy  841 CGCACGCCAGAGGTTGCCCGCGCCGCGCCACCCCGTTATTGGATTTGATCATGGCAGCG 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 CGCACGCCAGAGGTTGCCCGCGCCGCGCCACCCCGTTATTGGATTTGATCATGGCAGCG 900

Qy  901 TTGACGCCCCATCCACCCGCAAAAACAGGGGTATGGTGTGACATTACCCACTTCAGTACTG 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TTGACGCCCCATCCACCCGCAAAAACAGGGGTATGGTGTGACATTACCCACTTCAGTACTG 960

Qy  961 TTTATTGCCGGACACGATACTAATCTCGGCAAAATCTCGGCGGAGCTCAAACTGG 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 TTTATTGCCGGACACGATACTAATCTCGGCAAAATCTCGGCGGAGCTCAAACTGG 1020

Qy 1021 ACGCTTCCCGGTCAGCCGGATAACACGCCGCCCAGGTGGGTGAACTGGTTTGAACGCTGG 1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 ACGCTTCCCGGTCAGCCGGATAACACGCCGCCCAGGTGGGTGAACTGGTTTGAACGCTGG 1080

Qy 1081 CGTCGGGCTAAGCCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCAGACTTTACAG 1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 CGTCGGGCTAAGCCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCAGACTTTACAG 1140
```

RESULT 4

```
LOCUS       AR492966              1323 bp    DNA     linear   PAT 15-MAY-2004
DEFINITION  Sequence 1 from patent US 6720014.
ACCESSION   AR492966
VERSION     AR492966.1  GI:47264255
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1323)
  AUTHORS   Short,J.M. and Kretz,K.A.
  TITLE     Phytase-containing foodstuffs and methods of making and using them
  JOURNAL   Patent: US 6720014-A 1 13-APR-2004;
            Diversa Corporation; San Diego, CA
FEATURES             Location/Qualifiers
     source          1..1323
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match             100.0%;  Score 1322;  DB 2;  Length 1323;
  Best Local Similarity   100.0%;
  Matches 1323;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```
Qy          1 ATGAAAGCGATCTTAATCCCATTTTTATCTTTACTCTTCCGTTAACCCGCGAATCTGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAAGCGATCTTAATCCCATTTTTATCTTTACTCTTCCGTTAACCCGCGAATCTGCA 60

Qy         61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAAGTGTGGTGATGTGTCAGTCGTCAGGT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAAGTGTGGTGATGTGTCAGTCGTCAGGT 120

Qy        121 GTGCGTGCTCCAACCAAGGGCCACGCAACTGATGCAGGATGTCACCCAGACCGCATGGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTGCGTGCTCCAACCAAGGGCCACGCAACTGATGCAGGATGTCACCCAGACCGCATGGCCA 180

Qy        181 ACCTGGCGCGGTAAAACTGGGGTTGGCTGACACACCGCGACCGCGCGTAATCGCCTATCTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCTGGCGCGGTAAAACTGGGGTTGGCTGACACACCGCGACCGCGCGTAATCGCCTATCTC 240

Qy        241 GGACATTACCAACGCCAGCGTCTGGTAGCCGAGATTGCTCGGCGAAAAAGGGCTGCCCG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGACATTACCAACGCCAGCGTCTGGTAGCCGAGATTGCTCGGCGAAAAAGGGCTGCCCG 300

Qy        301 CAGTCTCGGTCAGGTCGGCGATTATTGCTGAATGTCGACGAGGGTACCCGTAAAACAGGCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGTCTCGGTCAGGTCGGCGATTATTGCTGAATGTCGACGAGGGTACCCGTAAAACAGGCGGA 360

Qy        361 GCCTTCGCCGCCGCGGCTGGCGACCTGACGTGTGCAATAACCGTAACATACCCAGATACG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTTCGCCGCCGCGGCTGGCGACCTGACGTGTGCAATAACCGTAACATACCCAGATACG 420

Qy        421 TCCAGTCCGATCCGTTATTTAATCCTCTAAAAAACTGGCGTTTGCCAACTGGATAACGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCAGTCCGATCCGTTATTTAATCCTCTAAAAAACTGGCGTTTGCCAACTGGATAACGCG 480
```

(... alignment continues 481–1080 ...)

```
Qy       1141 CAGATGGCTGATAAAACCGGCGGATTGAAAGGTGAAGGTGAAACTGACC 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CAGATGGCTGATAAAACCGGCGGATTGAAAGGTGAAGGTGAAACTGACC 1200

Qy       1201 CTGGCCAGGACGGATGTGAAGACGGAATGCCAGGGCCATGTTTTACCGCAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CTGGCCAGGACGGATGTGAAGACGGAATGCCAGGGCCATGTTTTACCGCAA 1260

Qy       1261 ATCGTGAATGAAGCACGCGATACCGGCGGTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ATCGTGAATGAAGCACGCGATACCGGCGGTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320

Qy       1321 TAA 1323
              |||
Db       1321 TAA 1323
```

RESULT 5

```
LOCUS       AR636185              1323 bp    DNA     linear   PAT 20-APR-2005
DEFINITION  Sequence 1 from patent US 6855365.
ACCESSION   AR636185
VERSION     AR636185.1  GI:62768868
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
```

```
Qy        421 TCCAGTCCGATCCGTTATTTAATCCTCTAAAAAACTGGCGTTTGCCAACTGGATAACGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCAGTCCGATCCGTTATTTAATCCTCTAAAAAACTGGCGTTTGCCAACTGGATAACGCG 480

Qy        481 AACGTGACTGACCGCGATCCTCAGCAGCGGCAGGAGGTCAATTGCTGACTTTACCGGCGGCAT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACGTGACTGACCGCGATCCTCAGCAGCGGCAGGAGGTCAATTGCTGACTTTACCGGCGGCAT 540

Qy        541 CGGCAAACGGCGTTTCCGGAACTGGACACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGGCAAACGGCGTTTCCGGAACTGGACACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600

Qy        601 CTTAAACGTGAGAAACAGGAACGAAAGCTTGTTCATTAACGTCACCAGCCATTACCATGGGAACTC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTAAACGTGAGAAACAGGAACGAAAGCTTGTTCATTAACGTCACCAGCCATTACCATGGGAACTC 660

Qy        661 AAGGTGAGCGCCGACAATGTCTCATTAACGGTGCGGTTAAGCCTTCGCATCAATGCTGACG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGGTGAGCGCCGACAATGTCTCATTAACGGTGCGGTTAAGCCTTCGCATCAATGCTGACG 720

Qy        721 GAGATATTTCTCCTGCAACAACGACACAGGGAAATGCCGGAGCCGGGTGGGGAGGATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGATATTTCTCCTGCAACAACGACACAGGGAAATGCCGGAGCCGGGTGGGGAGGATCACC 780

Qy        781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTATTGCTACAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTATTGCTACAA 840

Qy        841 CGCACGCCAGAGGTTGCCCGCACCGCCACCCCGTTATTGGATTGATCATGGCACGCG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CGCACGCCAGAGGTTGCCCGCACCGCCACCCCGTTATTGGATTGATCATGGCACGCG 900

Qy        901 TTGACGCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTACTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTGACGCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTACTG 960

Qy        961 TTTATTGCCGGACACGATTACTAATCTGGCAAAATCTCGGCGGCGCCACTGGGAGCTCAACTGG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTTATTGCCGGACACGATTACTAATCTGGCAAAATCTCGGCGGCGCCACTGGGAGCTCAACTGG 1020

Qy       1021 ACGCTTCCCGGTCAGCCGGATAACACGCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACGCTTCCCGGTCAGCCGGATAACACGCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1080

Qy       1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGGCTGGTCTTCCAGACTTTACAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGGCTGGTCTTCCAGACTTTACAG 1140

Qy       1141 CAGATGGCTGATAAAACCGGCGGATTGAAAGGTGAAGGTGAAACTGACC 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CAGATGGCTGATAAAACCGGCGGATTGAAAGGTGAAGGTGAAACTGACC 1200

Qy       1201 CTGGCCAGGACGGATGTGAAGACGGAATGCCAGGGCCATGTTTTACCGCAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CTGGCCAGGACGGATGTGAAGACGGAATGCCAGGGCCATGTTTTACCGCAA 1260

Qy       1261 ATCGTGAATGAAGCACGCGATACCGGCGGTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ATCGTGAATGAAGCACGCGATACCGGCGGTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320

Qy       1321 TAA 1323
              |||
Db       1321 TAA 1323
```

REFERENCE
AUTHORS   Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B.,
           O'Donoghue,B. and Mathur,E.J.,
TITLE   Recombinant bacterial phytases and uses thereof
JOURNAL   Patent: US 6855365-A 1 15-FEB-2005;
           Diversa Corporation; San Diego, CA
FEATURES   Location/Qualifiers
  source   1..1323
                  /organism="unknown"
                  /mol_type="genomic DNA"
ORIGIN

Query Match   100.0%; Score 1322; DB 2; Length 1323;
Best Local Similarity   100.0%; Pred. No. 0; Mismatches 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60

Qy     61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTCGTGATCGTCAGTCGTCATGGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTCGTGATCGTCAGTCGTCATGGT 120

Qy    121 GTGCGTGCTCCAACCAAGGCCACGCAGCTGATGCAGGATGTCACCCCAGACGCATGGCCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGCGTGCTCCAACCAAGGCCACGCAGCTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy    181 ACCTGGCCGGGTAAAACTGGGTTGGCTGACACCGCGCGGTGGTGAGCTAATCGCCTATCTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACCTGGCCGGGTAAAACTGGGTTGGCTGACACCGCGCGGTGGTGAGCTAATCGCCTATCTC 240

Qy    241 GGACATTACCAACGCCAGCGTCTCGGTAGCCGACGGATTGCTGCGCGAAAAAGGCAGGGAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGACATTACCAACGCCAGCGTCTCGGTAGCCGACGGATTGCTGCGCGAAAAAGGCAGGGAA 300

Qy    301 CAGTCTCGGTTCAGTGGAATTATTGCTGATGTCGACGAGCGTACCCGTTAAACAGGCGAA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGTCTCGGTTCAGTGGAATTATTGCTGATGTCGACGAGCGTACCCGTTAAACAGGCGAA 360

Qy    361 GCCTTCGCCGCGCGGGGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCTTCGCCGCGCGGGGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420

Qy    421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAAACGCG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAAACGCG 480

Qy    481 AACGTGACTGACCGGATCCTCAGCAGGGCGAGAGGGTCAATTCCTGACTTACTGCGGCAT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AACGTGACTGACCGGATCCTCAGCAGGGCGAGAGGGTCAATTCCTGACTTACTGCGGCAT 540

Qy    541 CGGCAAACGGCGGTTTCGGGAACGGAGGTCGCTTAATTTCCGGCAATCAAACTTTGTGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CGGCAAACGGCGGTTTCGGGAACGGAGGTCGCTTAATTTCCGGCAATCAAACTTTGTGC 600

Qy    601 CTTAAACGTGAGAAACAGGACAGAAAGTGTTCATTACGCAGCGCATTACCATCTGGAACTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTTAAACGTGAGAAACAGGACAGAAAGTGTTCATTACGCAGCGCATTACCATCTGGAACTC 660

Qy    661 AAGGTGAGCGCGCCGACGCGAATGCTCTCATTAACCGGTGCGGTGAAGCTCGCATCAATGCTGACG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AAGGTGAGCGCGCCGACGCGAATGCTCTCATTAACCGGTGCGGTGAAGCTCGCATCAATGCTGACG 720

Qy    721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGGTGGGGAGGGATCACC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGGTGGGGAGGGATCACC 780

Qy    781 GATTCACCACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAAATTTATTATTGCTACAA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GATTCACCACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAAATTTATTATTGCTACAA 840

Qy    841 CGCCACGCCAGAGAGGTTGCCCGCAGCCCGCGGGCCACCCGGTTATTGGATTTGATCATGGCAGCG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CGCCACGCCAGAGAGGTTGCCCGCAGCCCGCGGGCCACCCGGTTATTGGATTTGATCATGGCAGCG 900

Qy    901 TTGACGGCGCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACCTTCAGTACTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTGACGGCGCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACCTTCAGTACTG 960

Qy    961 TTTATTGCCGGACACGGATACTAATCTGGCAAAATCTCGGCGCGCGCGGCCGCACTGGAGCTCAACTGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTTATTGCCGGACACGGATACTAATCTGGCAAAATCTCGGCGCGCGCGGCCGCACTGGAGCTCAACTGG 1020

Qy   1021 ACGCTTCCCGGTCAGCCGGATAAACACGCCCGCCAGTGGTGAACTGGTGTTGAACGGCTGG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGCTTCCCGGTCAGCCGGATAAACACGCCCGCCAGTGGTGAACTGGTGTTGAACGGCTGG 1080

Qy   1081 CGTCGGCTAAGGCGGATPAACACGCCCAGTGGGATTCAGGTTTCAGGTTCCCAGACTTTACAG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CGTCGGCTAAGGCGGATPAACACGCCCAGTGGGATTCAGGTTTCAGGTTCCCAGACTTTACAG 1140

Qy   1141 CAGATGCGTGATAAAAACGCGCGTCATTAAATACGGCCGGCGAGGTGAAAACTGACC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAGATGCGTGATAAAAACGCGCGTCATTAAATACGGCCGGCGAGGTGAAAACTGACC 1200

Qy   1201 CTGGCAGGATGTGAAGAGCGGAAAATGCCGCAGGGCCATGTGGTCAGTTGGTGGCAGTTTACGGCAA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGCAGGATGTGAAGAGCGGAAAATGCCGCAGGGCCATGTGGTCAGTTGGTGGCAGTTTACGGCAA 1260

Qy   1261 ATCGTGAATGAAAGCACGCATTACCCGCGTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAAGCACGCATTACCCGCGTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320

Qy   1321 TAA 1323
          |||
Db   1321 TAA 1323
```

RESULT 6

LOCUS   AX052713   1323 bp   DNA   linear   PAT 12-JAN-2001
DEFINITION Sequence 1 from Patent WO0071728.
ACCESSION   AX052713
VERSION   AX052713.1 GI:12226903
KEYWORDS   .
SOURCE   Escherichia coli
  ORGANISM Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Short,J.M. and Kretz,K.A.
  TITLE   Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: WO 0071728-A 1 30-NOV-2000;
            Diversa Corporation (US)
FEATURES   Location/Qualifiers
     source   1..1323
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
     CDS   1..1323
                     /codon_start=1
                     /transl_table=11
                     /note="unnamed protein product"
                     /protein_id="CAC21810.1"
                     /db_xref="GI:12226904"
                     /translation="MKAILIPFLSLLIPLTFQSSAFAQSEPELKLESVVIVSRHGVRAP
                     TKATQLMQDVTPDAWPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLALAKKGCPQS
                     GQVAIIADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA
                     NVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPS
                     ELKVSADNVSITGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQF
                     YLLQRTPEVARSRATPLLDLIMAALTPHPPDKQAYGVTLPTSVLFIAGHDTNLANLGG
                     ALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNT
                     PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSLRSHHHHHH"
ORIGIN

```
Query Match          100.0%;  Score 1322;  DB 2;  Length 1323;
Best Local Similarity 100.0%;  Pred. No. 0;  Mismatches 0;
Matches 1323;  Conservative 0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1 ATGAAAGCGGATCTTAATCCATTTTTATCTCTTCTTGATTCCGTTAACCCCGCCGAAATCTGCA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAAAGCGGATCTTAATCCATTTTTATCTCTTCTTGATTCCGTTAACCCCGCCGAAATCTGCA 60

Qy    61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGTCGGTGATTGTCAGTCGTCATGGT 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGTCGGTGATTGTCAGTCGTCATGGT 120

Qy   121 GTGCGTGCTCCAACCAAGGCCACCGCAACTGATGCAGGATGTCACCCCAGAGGCCATGGCCA 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTGCGTGCTCCAACCAAGGCCACCGCAACTGATGCAGGATGTCACCCCAGAGGCCATGGCCA 180

Qy   181 ACCTGGCCGGGTAAAACTGGGTTGGCGTTGACACGCGGCTAATCCCTATCTC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ACCTGGCCGGGTAAAACTGGGTTGGCGTTGACACGCGGCTAATCCCTATCTC 240

Qy   241 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGTGTTGCTGGCGGAAAAGGGGCTGCCCG 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGTGTTGCTGGCGGAAAAGGGGCTGCCCG 300

Qy   301 CAGTCTCGGTCAGGTCCGGATCCTCAGCGCAGGGGTCAAATTGCTGACTTTTACCGGGCAT 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGTCTCGGTCAGGTCCGGATCCTCAGCGCAGGGGTCAAATTGCTGACTTTTACCGGGCAT 360

Qy   361 GCCTTCGCCTGCCGCGGGTGGCACCTGACTGTGCAATAACCGTACATATCCCAGATACG 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCCTTCGCCTGCCGCGGGTGGCACCTGACTGTGCAATAACCGTACATATCCCAGATACG 420

Qy   421 TCCAGTCCGATCCGTTATTTAATCCTCTAAAAACTGGATAACGGCG 480
         |||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCCAGTCCGATCCGTTATTTAATCCTCTAAAAACTGGATAACGGCG 480

Qy   481 AACGTGACTGAGGCGATCCTCAGCAGGGGTCAAATTGCTGACTTTACCGGGCAT 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AACGTGACTGAGGCGATCCTCAGCAGGGGTCAAATTGCTGACTTTACCGGGCAT 540

Qy   541 CGGCAAACGGGCGGTTTCCGGACTGGGAACGGCGTGCTTAATTTTCCGGAATCAAACTTGTGC 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CGGCAAACGGGCGGTTTCCGGACTGGGAACGGCGTGCTTAATTTTCCGGAATCAAACTTGTGC 600

Qy   601 CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCATTAACGCAGCATTACCATCGGAACTC 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCATTAACGCAGCATTACCATCGGAACTC 660

Qy   661 AAGGTGAGCGCCGACAAATGTCTCATTAACCCGTGCGCGTAAGCCTCCGCCAATGCTGACG 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AAGGTGAGCGCCGACAAATGTCTCATTAACCCGTGCGCGTAAGCCTCCGCCAATGCTGACG 720

Qy   721 GAGATATTTCTCCTGCAACAAGCACAGGGATCACC 780
         |||||||||||||||||||||||||||||||||||
Db   721 GAGATATTTCTCCTGCAACAAGCACAGGGATCACC 780

Qy   781 GATTCACACAGTGGAACACCTTGCTAAGTTTGCATAAGCGGCCAATTTTATTTATTGCTACACA 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GATTCACACAGTGGAACACCTTGCTAAGTTTGCATAAGCGGCCAATTTTATTTATTGCTACACA 840

Qy   841 CGCACCGCCGAGGGTTGCCCGCCGCCACCCCGTTATTGGATTTGATCATTGGCAGCG 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CGCACCGCCGAGGGTTGCCCGCCGCCACCCCGTTATTGGATTTGATCATTGGCAGCG 900

Qy   901 TTGACCGCCCCATCCACCGCAAAAAACAGGCGTATGGGTGTGACATTACCCACTTCAGTACTG 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTGACCGCCCCATCCACCGCAAAAAACAGGCGTATGGGTGTGACATTACCCACTTCAGTACTG 960

Qy   961 TTTATTGCCGGACACGGATACTAATCTCGGCAAAATCTCGGCGCCACTGGAGCTCAACTGG 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TTTATTGCCGGACACGGATACTAATCTCGGCAAAATCTCGGCGCCACTGGAGCTCAACTGG 1020

Qy  1021 ACGCTTCCCGGTCAGGTCGGTGAACACGCGCCGGATAACACGCCGGCCAGGTCGGTGAACTGGTTCTTCCAGACTTTACAG 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 ACGCTTCCCGGTCAGGTCGGTGAACACGCGCCGGATAACACGCCGGCCAGGTCGGTGAACTGGTTCTTCCAGACTTTACAG 1080

Qy  1081 CGTCGGCTTAAGCCGATAACAGCACAGCACCAGTGGATTCAGTTTCGCTTCGTCTTCCAGAC 1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 CGTCGGCTTAAGCCGATAACAGCACAGCACCAGTGGATTCAGTTTCGCTTCGTCTTCCAGAC 1140

Qy  1141 CAGATGCGTGATGAAAAACGGCCGTCTGTTCATTAAATAACGGCG 1200
         |||||||||||||||||||||||||||||||||||||||||||||
Db  1141 CAGATGCGTGATGAAAAACGGCCGTCTGTTCATTAAATAACGGCG 1200

Qy  1201 CTGGCAGGATGTGAAGAGGCGAAAATGCCGCAGGGGCCATGGTTGTTCGTTGGCAGGTTTTTACGGCAA 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 CTGGCAGGATGTGAAGAGGCGAAAATGCCGCAGGGGCCATGGTTGTTCGTTGGCAGGTTTTTACGGCAA 1260

Qy  1261 ATCGTGAATGAAGCACGGCATACCGGCGCGCAGTTTGAGATCTCATCACCATCACCAC 1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 ATCGTGAATGAAGCACGGCATACCGGCGCGCAGTTTGAGATCTCATCACCATCACCAC 1320

Qy  1321 TAA 1323
         |||
Db  1321 TAA 1323
```

RESULT 7

```
LOCUS       AX338966                1323 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 1 from Patent WO0189317.
ACCESSION   AX338966
VERSION     AX338966.1  GI:18129102
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Short,J.M., Kretz,K.A. and O'Donoghue,E.
  TITLE     Dietary aids and methods of use thereof
  JOURNAL   Patent: WO 0189317-A 1 29-NOV-2001;
            DIVERSA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..1323
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
     CDS             1..1323
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAD20553.1"
                     /db_xref="GI:18129103"
                     /translation="MKAILIPFLSLLIPLTPQSAPAQSEPELKLESVVISVRHGVRAP
                     TKATQLMQDVTPDAWPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGGCPQS
                     GQVAIIADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDLPLNFPLKTGVCQLDNA
                     NVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPS
                     ELKVSADNVSLTGAVSLASMLTEIPLLQQAQGMPEPGWGRITDSHOWNTLLSLHNAQF
                     YLLQRTPEVARSRATFLLDLIMAALTPHPPDQKOAYGVTLPTSVLFIAGHDTNLANLGG
                     ALELNWTLPGQPDNTPPGGELVFERWRRLSDNSOWIQVSLVFQTLQQMRDKTPLSLNT
                     PPGEVKLITLAGCEERNAQGMCSLAGFTQIVNEARIPACSLRSHHHHHH"
ORIGIN
```

Query Match   100.0%; Score 1322; DB 2; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1 ATGAAAGCGGATCTTAATCCATTTTTATCTCTTCTTGATTCCGTTAACCCCGCCGAAATCTGCA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAAAGCGGATCTTAATCCATTTTTATCTCTTCTTGATTCCGTTAACCCCGCCGAAATCTGCA 60

Qy    61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGTCGGTGATTGTCAGTCGTCATGGT 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGTCGGTGATTGTCAGTCGTCATGGT 120
```

```
Qy    121 GTGCGTGCTTCCACCAAGGCCACGGCACTGATGCAGGAGGAGTCACCCCCAGACGCATGGGCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGCGTGCTTCCACCAAGGCCACGGCACTGATGCAGGAGGAGTCACCCCCAGACGCATGGGCA 180

Qy    181 ACCTGGCGCCGGTAAAACTGGGTTGGCTGACACCGCGGCGGGNGCGGTGGGTGGAGCTAATCGCTTATCTC 240
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db    181 ACCTGGCGCCGGTAAAACTGGGTTGGCTGACACCGCGG    GTGGGTGGAGCTAATCGCTTATCTC 240

Qy    241 GGACATTACCAACCCAGGCTCTGGTAGCCGACGGATTGCCTGGGGAAAAAGGGGCTGCCCG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGACATTACCAACCCAGGCTCTGGTAGCCGACGGATTGCCTGGGGAAAAAGGGGCTGCCCG 300

Qy    301 CAGTCTGGTTCAGGTCAGGTTCGGCGATTATTGTCGATGTCGACGAGCGTACCCGTAAAAACAGGGCGAA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGTCTGGTTCAGGTCAGGTTCGGCGATTATTGTCGATGTCGACGAGCGTACCCGTAAAAACAGGGCGAA 360

Qy    361 GCCTTCGCGCCGCGCTGGCACTGGCTGGCTGACCCGGCGGGNGGCTAATGCCTTATCTC 420
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCTTCGCGCCGCGCTGGCACTGGCTGACCCGGCGGG    CTAATGCCTTATCTC 420

Qy    421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAAACTGGGATGGTAACCGCG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAAACTGGGATGGTAACCGCG 480

Qy    481 AACGTGACTGACCCGGATCCTCACGCAGGCGCGAGGAGGGTCAAATTGCTGAGCGTGACATTTACCGGGCAT 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AACGTGACTGACCCGGATCCTCAGCAGGGCGGAGGGTCAAATTGCTGACTTTACCGGGCAT 540

Qy    541 CGGCAAACGGCGCTTTCGGCGAACTGGAACGGGTTGCTTAATTTCCGCAATCAAACTTGTGC 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CGGCAAACGGCGTTTCGGCGAACTGGAACGGGTTGCTTAATTTCCGCAATCAAACTTGTGC 600

Qy    601 CTTAAACGTGAGAAACAGGAGCGAAAGCTGTTCATTAACGGACCATTACCATCTGGAACTC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTTAAACGTGAGAAACAGGAGCGAAAGCTGTTCATTAACGGACCATTACCATCTGGAACTC 660

Qy    661 AAGGTGAGCGCCGACCAATGTCTCATTAACCGGTGCGGTTGGGGTAAGCCTCCGCATCAATGCTCACG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AAGGTGAGCGCCGACCAATGTCTCATTAACCGGTGCGGTTAAGCCTCCGCATCAATGCTCACG 720

Qy    721 GAGATATTTCTCCTGCAACAAGCAATCGGGAGCGGGTGGGGAGGGATCACC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCCGGACGGTGGGGAGGGATCACC 780

Qy    781 GATTCACACCAGTGGGAACACCTTGCTAAGTTTGCATAACGGCGCAATTTTATTGCTACAA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GATTCACACCAGTGGGAACACCTTGCTAAGTTTGCATAACGGCGCAATTTTATTGCTACAA 840

Qy    841 CGCACGCCAGAGGTTGCCCGGCCAGCCCGGCCACCCGGTTATTGGATTTGGATTCATGGCCAGCG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CGCACGCCAGAGGTTGCCCGGCCAGCCCGGCCACCCGGTTATTGGATTTGGATTCATGGCCAGCG 900

Qy    901 TTGACGCCCCATCCACCGCAAAAACAGGGCGTATGGGTGACATTACCACTTCAGTACTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTGACGCCCCATCCACCGCAAAAACAGGGCGTATGGGTGACATTACCACTTCAGTACTG 960

Qy    961 TTTATTCGCGGACGACTAATCTGCGAAATCCGCGGGCGGCCCAGCTGAGCGTTCAACTGG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTTATTCGCGGACGACTAATCTGCGAAATCCGCGGGCGGCCCAGCTGAGCGTTCAACTGG 1020

Qy   1021 ACGCTTCCCGGTCAGCGGGATAACACGCGCCGGATAACTAATCTGGCAAATCTGGGG 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGCTTCCCGGTCAGCGGGATAACACGCGCCGGATAACTAATCTGGCAAATCTGGGG 1080

Qy   1081 CGTCGGCTAAGCGGATAACAGCCAGTGGAATCAGGTTTCGCTGGTTTGCTGGTGTCTTCCAGACTTTACAG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CGTCGGCTAAGCGGATAACAGCCAGTGGAATCAGGTTTCGCTGGTTTGCTGGTGTCTTCCAGACTTTACAG 1140

Qy   1141 CAGATGGCTGATAAAATACGGCCGCCGCCGCTGCTGTCATTAAATACGGCTGACC 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAGATGGCTGATAAAATACGGCCGCCGCCGCTGCTGTCATTAAATACGGCTGACC 1200

Qy   1201 CTGGCAGGATGTGAAGAGCGAAATGCCCAGGGGCCGAAATGCGGCAGGGCATGTGTTCGTTGTTGGCAGGTTTACGGCAA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGCAGGATGTGAAGAGCGAAATGCCCAGGGGCCGAAATGCGGCAGGGCATGTGTTCGTTGTTGGCAGGTTTACGGCAA 1260

Db   1261 ATCCGTGAATGAAGCACGCATTACCGCGTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320

Qy   1261 ATCCGTGAATGAAGCACGCATTACCGCGTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320

Db   1321 TAA 1323

Qy   1321 TAA 1323
```

RESULT 8

```
LOCUS       AX356566                1323 bp    DNA        linear    PAT 06-FEB-2002
DEFINITION  Sequence 1 from Patent WO0190333.
ACCESSION   AX356566
VERSION     AX356566.1  GI:18621053
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B. and
            O'Donoghue,B.
  TITLE     Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: WO 0190333-A 1 29-NOV-2001;
            DIVERSA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..1323
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
     CDS             1..1323
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAD23025.1"
                     /db_xref="GI:18621054"
                     /translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAP
                     TKATQLMQDVTPDAWPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQS
                     GQVAIIADVDERRTKIGEAFAAGLAPDCAITVHTQADTSSPDDLFNPLKTGVCOLDNA
                     NVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKRKKQDBESCSLITQALPS
                     ELKVSADNVSLTGAVSLASMLTEIFLLQQAQMPEPGWGRITDSHQWNTLLSLHNAQF
                     YLLQRPTEVARSRATPLLDLIMAALTPHPPQQKOAYGVTLPTSVLFIAGHDTNIANLGG
                     ALBLANWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTFPLSLNT
                     PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSLRSHHHHHH"
ORIGIN
```

```
Query Match        100.0%;   Score 1322;  DB 2;  Length 1323;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 1323;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTTCTTCGGTTAACCCCGCAATCTGCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTTCTTCGGTTAACCCCGCAATCTGCA 60

Qy     61 TTCGCTCAGAGTGAGCCCGAGCTGAAGCTGGAAAGTGTGGTCGGTCACCGGCATGGTCGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCGCTCAGAGTGAGCCCGAGCTGAAGCTGGAAAGTGTGGTCGGTCACCGGCATGGTCGT 120

Qy    121 GTGCGTGCTCCACCAAGGCCACGGCACTGATGCAGGACGTCACCCCAGACGCATGGGCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGCGTGCTCCACCAAGGCCACGGCACTGATGCAGGACGTCACCCCAGACGCATGGGCA 180

Qy    181 ACCTGGCCCGGTAAAACTGGGTTGGCTGACACCGCGTGGTGGTGAGCTGGCTTACTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACCTGGCCCGGTAAAACTGGGTTGGCTGACACCGCGTGGTGGTGAGCTGGCTTACTC 240

Qy    241 GGACATTACCAACCGCACGCCAGCGTCTGGTAGCCGACGGATTGCTGGCCGAAAAAAGGGCTGCCCG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGACATTACCAACCGCACGCCAGCGTCTGGTAGCCGACGGATTGCTGGCCGAAAAAAGGGCTGCCCG 300
```

RESULT 9

```
LOCUS       AR636190                1901 bp    DNA     linear   PAT 20-APR-2005
DEFINITION  AR636190 Sequence 7 from patent US 6855365.
ACCESSION   AR636190
VERSION     AR636190.1  GI:62768873
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1901)
  AUTHORS   Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B.,
            O'Donoghue,B. and Mathur,E.J.
  TITLE     Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: US 6855365-A 7 15-FEB-2005;
            Diversa Corporation; San Diego, CA
FEATURES             Location/Qualifiers
     source          1..1901
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN
```

Query Match   97.4%; Score 1287.8; DB 2; Length 1901;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1293;   Conservative   0; Mismatches   7; Indels   0; Gaps   0;

```
RESULT 10
LOCUS       AX356572                1901 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 7 from Patent WO0190333.
ACCESSION   AX356572
VERSION     AX356572.1  GI:18621059
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B. and
            O'Donoghue,E.
  TITLE     Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: WO 0190333-A 7 29-NOV-2001;
            DIVERSA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..1901
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

Query Match     97.4%;  Score 1287.8;  DB 2;  Length 1901;
Best Local Similarity   99.5%;
Matches 1293;  Conservative   0;  Mismatches   7;  Indels   0;  Gaps   0;
```

(Sequence alignment — Query (Qy) versus Database (Db), paired rows with position numbers and identity bars.)

```
Qy  1141  CAGATCGTGATAAAACGGCGCTGTGCTGTCATTAAAATACGGCGCCCGGAGGAGGTGAAACTGACC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1328  CAGATCGTGATAAAACGGCGCTGTGCTGTCATTAAAATACGGCGCCCGGAGGAGGTGAAACTGACC 1387

Qy  1201  CTGGCAGGATGTGAAGAGCGAAATGCCCAGGGCATGTGTTCGTTGGCGCAGGTTTTACCCAA 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1388  CTGGCAGGATGTGAAGAGCGAAATGCCCAGGGCATGTGTTCGTTGGCGCAGGTTTTACCCAA 1447

Qy  1261  ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTGAGAT 1300
          ||||||||||||||||||||||||||||||||||||||
Db  1448  ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTGTAAT 1487
```

```
RESULT 11
LOCUS       ECOAPPAA
DEFINITION  Escherichia coli periplasmic phosphoanhydride phosphohydrolase
            (appA) gene, complete cds.
ACCESSION   M58708
VERSION     M58708.1  GI:145283
KEYWORDS    .
SOURCE      Escherichia coli
ORGANISM    Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1  (bases 1 to 1901)
  AUTHORS   Dassa,J., Marck,C. and Boquet,P.L.
  TITLE     The complete nucleotide sequence of the Escherichia coli gene appA
            reveals significant homology between pH 2.5 acid phosphatase and
            glucose-1-phosphatase
  JOURNAL   J. Bacteriol. 172 (9), 5497-5500 (1990)
   PUBMED   2168385
REFERENCE   2  (sites)
  AUTHORS   Ostanin,K., Harms,E.H., Stevis,P.E., Kuciel,R., Zhou,M.M. and Van
            Etten,R.L.
  TITLE     Overexpression, site-directed mutagenesis, and mechanism of
            Escherichia coli acid phosphatase
  JOURNAL   J. Biol. Chem. 267 (32), 22830-22836 (1992)
   PUBMED   1429631
REFERENCE   3  (sites)
  AUTHORS   Ostanin,K. and Van Etten,R.L.
  TITLE     Asp304 of Escherichia coli acid phosphatase is involved in leaving
            group protonation
  JOURNAL   J. Biol. Chem. 268 (28), 20778-20784 (1993)
   PUBMED   8407904
```

```
ECOAPPAA        1901 bp    DNA    linear    BCT 04-APR-2002
```

```
FEATURES             Location/Qualifiers
     source          1..1901
                     /organism="Escherichia coli"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:562"
     gene            3..1522
                     /gene="appA"
     RBS             3..8
                     /gene="appA"
     CDS             16..108
                     /gene="appA"
                     /note="ORFX"
                     /codon_start=1
                     /transl_table=1
                     /product="unknown"
                     /protein_id="AAA72085.1"
                     /db_xref="GI:145284"
                     /translation="MWYLLWFVGILLMCSLSTLVLVWLDPRLKS"
     stem_loop       118..149
                     /gene="appA"
     -35_signal      133..138
                     /gene="appA"
     -10_signal      156..161
                     /gene="appA"
     RBS             172..177
                     /gene="appA"
     CDS             188..1486
                     /gene="appA"
                     /standard_name="acid phosphatase"
                     /codon_start=1
                     /transl_table=11
                     /product="periplasmic phosphoanhydride phosphohydrolase"
                     /protein_id="AAA72086.1"
                     /db_xref="GI:145285"
                     /translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP
                     TKATQLMQDVTPDAWPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQS
                     GQVAIIADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA
                     NVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPS
                     ELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQF
                     YLLQRTPEVARSRATPLLDLIKTALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANIGG
                     ALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNT
                     PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL"
     sig_peptide     188..253
                     /gene="appA"
     mat_peptide     254..1483
                     /gene="appA"
                     /product="periplasmic phosphohydride phosphohydrolase"
     variation       299..301
                     /gene="appA"
                     /standard_name="H303A mutant"
                     /note="created by site directed mutagenesis"
                     /citation=[2]
     variation       299..301
                     /gene="appA"
                     /standard_name="R16A mutant"
                     /note="created by site directed mutagenesis"
                     /citation=[2]
     variation       299..301
                     /gene="appA"
                     /standard_name="R20A mutant"
                     /note="created by site directed mutagenesis"
                     /citation=[2]
     variation       299..301
                     /gene="appA"
                     /standard_name="R92A mutant"
                     /note="created by site directed mutagenesis"
                     /citation=[2]
     variation       302..304
                     /gene="appA"
                     /standard_name="H17N mutant"
                     /note="created by site directed mutagenesis"
                     /citation=[2]
     variation       1163..1165
                     /gene="appA"
                     /standard_name="D304A mutant"
                     /note="created by site directed mutagenesis"
                     /citation=[3]
                     /phenotype="Asp 304 changed to Ala"
                     /replace="gct"
     variation       1163..1165
                     /gene="appA"
                     /standard_name="D304E mutant"
                     /note="created by site directed mutagenesis"
                     /citation=[3]
                     /phenotype="Asp 304 changed to Glu"
                     /replace="gag"
     terminator      1497..1522
                     /gene="appA"
                     /note="rho-independent terminator"
ORIGIN
```

```
Qy Query Match      97.4%;   Score 1287.8;   DB 15;   Length 1901;
Best Local Similarity   99.5%;   Pred. No. 0;
Matches 1293;   Conservative 0;   Mismatches 7;   Indels 0;   Gaps 0;

Qy   1  ATGAAAGCGAATCTTAATCCCATTTTATCTCTTCGATTCCGTTAACCCCGCGAATCTGCA 60
        |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  188  ATGAAAGCGAATCTTAATCCCATTTTATCTCTTCGATTCCGTTAACCCCGCGAATCTGCA 247

Qy  61  TTCGCTCAGAGTGAGCCGGAGCTGAAGACTGAAAAGTGTGGTTGGTGAATTGTCAGTCGTCATGGT 120
```

(Sequence alignment block — Query (Qy) vs Database (Db) pairwise alignment, with vertical bars indicating matches)

```
Db   248  ................................................  307
Qy   121  GTGCGGTGCTTCCA................................... 180
Db   308  GTGCGGTGCTTCCA................................... 367
Qy   181  ACCTTGGCGCCGGT.................................  240
Db   368  ACCTTGGGCGCCGGT................................  427
Qy   241  GGACATTACCAACC................................  300
Db   428  GGACATTACCAACC................................  487
Qy   301  CAGTCTGTTCAGGT................................  360
Db   488  CAGTCTGTTCAGGT................................  547
Qy   361  GCCTTTCGCGCCGC................................  420
Db   548  GCCTTCGCGCCGC.................................  607
Qy   421  TCCAGTCCCGATCC................................  480
Db   608  TCCAGTCCCGATCC................................  667
Qy   481  AAGGTGACTGACGC................................  540
Db   668  AAGGTGACTGACGC................................  727
Qy   541  CGGCAAACGGCGTT................................  600
Db   728  CGGCAAACGGCGTT................................  787
Qy   601  CTTAAACGTGAGAA................................  660
Db   788  CTTAAACGTGAGAA................................  847
Qy   661  AAGGTGAGCGCGCGGA..............................  720
Db   848  AAGGTGACCGCGCGGA..............................  907
Qy   721  GAGATATTTCTCCTG...............................  780
Db   908  GAGATATTTCTCCTG...............................  967
Qy   781  GATTCACACCAGTGG...............................  840
Db   968  GATTCACACCAGTGG...............................  1027
Qy   841  CGCACGCCCAGAGAG...............................  900
Db  1028  CGCACGCCCAGAGAG...............................  1087
Qy   901  TTGACGCCCCATCCA...............................  960
Db  1088  TTGACGCCCCATCCA...............................  1147
Qy   961  TTTATTGCCGGACAC...............................  1020
Db  1148  TTTATCGCCGGACAC...............................  1207
Qy  1021  ACGCTTCCCGGTCAG...............................  1080
Db  1208  ACGCTTCCCGGTCAG...............................  1267
Qy  1081  CGTCGGCTAAGCGGA...............................  1140
Db  1268  CGTCGGCTAAGCGGA...............................  1327
Qy  1141  CAGATGCCGTGATAAAACGCCGTCGTCATTAAATACGCCGCCGGAGAGGGTTGAAACTGACC  1200
Db  1328  CAGATGCCGTGATAAAACGCCGTCGTCATTAAATACGCCGCCGGAGAGGGTTGAAACTGACC  1387
Qy  1201  CTGGCAGGATGTGAAGAGCGAAAATCCGCACGGGCATGTGTTCGTTGCGTTCGGCAGTTTTACGGCAA  1260
Db  1388  CTGGCAGGATGTGAAGAGCGAAAATCCGCACGGGCATGTGTTCGTTGCGTTGGGCAGTTTTACGGCAA  1447
Qy  1261  ATCGTGAATGAAGCACGCAATACCGGGCGTGCAGTTTGTAAT  1300
Db  1448  ATCGTGAATGAAGCACGCAATACCGGGCGTGCAGTTTGTAAT  1487
```

RESULT 12
U00096_10
WPCOMMENT
WPCOMMENT
Sequence split into 47 fragments LOCUS U00096 Accession U00096

Fragment Name	Begin	End
U00096_00	1	110000
U00096_01	100001	210000
U00096_02	200001	310000
U00096_03	300001	410000
U00096_04	400001	510000
U00096_05	500001	610000
U00096_06	600001	710000
U00096_07	700001	810000
U00096_08	800001	910000
U00096_09	900001	1010000
U00096_10	1000001	1110000
U00096_11	1100001	1210000
U00096_12	1200001	1310000
U00096_13	1300001	1410000
U00096_14	1400001	1510000
U00096_15	1500001	1610000
U00096_16	1600001	1710000
U00096_17	1700001	1810000
U00096_18	1800001	1910000
U00096_19	1900001	2010000
U00096_20	2000001	2110000
U00096_21	2100001	2210000
U00096_22	2200001	2310000
U00096_23	2300001	2410000
U00096_24	2400001	2510000
U00096_25	2500001	2610000
U00096_26	2600001	2710000
U00096_27	2700001	2810000
U00096_28	2800001	2910000
U00096_29	2900001	3010000
U00096_30	3000001	3110000
U00096_31	3100001	3210000
U00096_32	3200001	3310000
U00096_33	3300001	3410000
U00096_34	3400001	3510000
U00096_35	3500001	3610000
U00096_36	3600001	3710000
U00096_37	3700001	3810000
U00096_38	3800001	3910000
U00096_39	3900001	4010000
U00096_40	4000001	4110000
U00096_41	4100001	4210000
U00096_42	4200001	4310000
U00096_43	4300001	4410000
U00096_44	4400001	4510000
U00096_45	4500001	4610000
U00096_46	4600001	4639675

Continuation (11 of 47) of U00096 from base 1000001 (U00096 Escherichia coli K-12 MG1655

Query Match 97.4%; Score 1287.8; DB 15; Length 110000;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1292; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy      1  ATGAAAGCGATCTTAATCCCATTTTATCTCTTCTGATTCGTTAACCCGCGAATCTGCA      60
Db  39840  ATGAAAGCGATCTTAATCCCATTTTATCTCTTCTGATTCGTTAACCCGCGAATCTGCA  39899
```

The left portion of the page contains a series of Qy/Db nucleotide sequence alignment blocks (query vs. database), each 60 bases per line with matching "|" indicators between them, running from Qy position 61 through 1200 and Db positions 39900 through 41100+.

RESULT 13
AP009048_10
WPCOMMENT
Sequence split into 47 fragments

Fragment Name   Begin   End   LOCUS AP009048 Accession AP009048
AP009048_00   1   110000
AP009048_01   100001   210000
AP009048_02   200001   310000
AP009048_03   300001   410000
AP009048_04   400001   510000
AP009048_05   500001   610000
AP009048_06   600001   710000
AP009048_07   700001   810000
AP009048_08   800001   910000
AP009048_09   900001   1010000
AP009048_10   1000001   1110000
AP009048_11   1100001   1210000
AP009048_12   1200001   1310000
AP009048_13   1300001   1410000
AP009048_14   1400001   1510000
AP009048_15   1500001   1610000
AP009048_16   1600001   1710000
AP009048_17   1700001   1810000
AP009048_18   1800001   1910000
AP009048_19   1900001   2010000
AP009048_20   2000001   2110000
AP009048_21   2100001   2210000
AP009048_22   2200001   2310000
AP009048_23   2300001   2410000
AP009048_24   2400001   2510000
AP009048_25   2500001   2610000
AP009048_26   2600001   2710000
AP009048_27   2700001   2810000
AP009048_28   2800001   2910000
AP009048_29   2900001   3010000
AP009048_30   3000001   3110000
AP009048_31   3100001   3210000
AP009048_32   3200001   3310000
AP009048_33   3300001   3410000
AP009048_34   3400001   3510000
AP009048_35   3500001   3610000
AP009048_36   3600001   3710000
AP009048_37   3700001   3810000
AP009048_38   3800001   3910000
AP009048_39   3900001   4010000
AP009048_40   4000001   4110000
AP009048_41   4100001   4210000
AP009048_42   4200001   4310000
AP009048_43   4300001   4410000
AP009048_44   4400001   4510000
AP009048_45   4500001   4610000
AP009048_46   4600001   4646332

Continuation (11 of 47) of AP009048 from base 1000001 (AP009048 Escherichia coli W3110 D

Query Match   97.4%;   Score 1287.8;   DB 15;   Length 110000;
Best Local Similarity 99.4%;   Pred. No: 0;   Mismatches 8;   Indels 0;   Gaps 0;
Matches 1292;   Conservative 0;   Mismatches 8;   Indels 0;   Gaps 0;

```
Qy      1 ATGAAAAGGGATCTTAATCCCATTTTTATCTCTTCTGATTCGGTTAACCCCGCAAATCTGCA 60
            |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  41039 ATGAAAGGGATCTTAATCCCATTTTTATCTCTTCTGATTCGGTTAACCCCGCAAATCTGCA 41098
```

```
RESULT 14
LOCUS       AR636188                1901 bp    DNA     linear   PAT 20-APR-2005
DEFINITION  Sequence 5 from patent US 6855365.
ACCESSION   AR636188
VERSION     AR636188.1  GI:62768871
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1901)
  AUTHORS   Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B.,
            O'Donoghue,E. and Mathur,E.J.
  TITLE     Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: US 6855365-A 5 15-FEB-2005;
            Diversa Corporation; San Diego, CA
FEATURES             Location/Qualifiers
     source          1..1901
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match             97.3%;  Score 1286.2;  DB 2;  Length 1901;
Best Local Similarity   99.4%;  Pred. No. 0;  Matches 8;  Indels 0;  Gaps 0;
Matches 1292;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;
```

RESULT 15

```
LOCUS       AR636191                1901 bp    DNA     linear   PAT 20-APR-2005
DEFINITION  Sequence 9 from patent US 6855365.
ACCESSION   AR636191
VERSION     AR636191.1  GI:62768874
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1901)
  AUTHORS   Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B.,
            O'Donoghue,E. and Mathur,E.J.
  TITLE     Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: US 6855365-A 9 15-FEB-2005;
            Diversa Corporation; San Diego, CA
FEATURES             Location/Qualifiers
     source          1..1901
                     /organism="unknown"
                     /mol_type="genomic DNA"
```

ORIGIN

```
Query Match             97.3%;  Score 1286.2;  DB 2;  Length 1901;
Best Local Similarity   99.4%;
Matches 1292;  Conservative    0;  Mismatches    0;  Indels    8;  Gaps    0;
```

```
Qy        1 ATGAAAGGGATCTTAATCCCATTTTTATCTCTTCGTTAACCCCGCAATCTGCA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      188 ATGAAAGGGATCTTAATCCCATTTTTATCTCTTCGTTAACCCCGCAATCTGCA 247

Qy       61 TTCGCTCAGAGTGAGCGGCGAGCTGAACCTTGGAAAAGTGTGGTGATTTGTCA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      248 TTCGCTCAGAGTGAGCCCGGAGCCTGGAAAGTGGTGATTTGTCAGTGCATGGT 307

Qy      121 GTGCGTCGTCCAACCAAGGCCAACGCAACTGATGCAGGATCACCCCAGACGGA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      308 GTGCGTGCTCCAACCAAGGCCAACGCAACTGATGCAGGATCACCCCAGACCGG 367

Qy      181 ACCTTGGCCGGTAAAACTGGGTTGGCTGACACCGGGGGTTGGGTGAGCTAATC 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      368 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGGGGGTTGGGTGAGCTAATCC 427

Qy      241 GGACATTACCCAACGGCCAGCGTCTTGGTAGCCGACGGATGCTGGGCGGATTG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      428 GGACATTACCCAACGGCCAGCGTCTTGGTAGCCGACGGATGCTGGGCGGATTG 487

Qy      301 CAGTCTCGGTCAGGTCGCGGATTATTAATCCTCTAAAAAACAGGCCAA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db      488 CAGTCTCGGTCAGGTCGCGGATTATTAATCCTCTAAAAAACAGGCCAA 547

Qy      361 GCCTTCGGCGCGCGGATTGGCACCTGACTGTGCAATAACCCAGACTACG 420
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db      548 GCCTTCGGCGCGCGGATTGGCACCTGACTGTGCAATAACCCAGACTACG 607

Qy      421 TCCAGTCCCGATCCGTTATTTTTAATCCCTCTAAAAAACTGGCGTTTTGCCAACTGATAACCG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      608 TCCAGTCCCGATCCGTTATTTTAATCCTCTAAAAAACTGGCGTTTTGCCAACTGATAACCG 667

Qy      481 AAGGTGACTGACGGCGATCCTCAGCAGGGGCAGAGGGTCAATTGCTGC 540
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db      668 AAGGTGACTGACGGCGATCCTCAGCAGGGGCAGAGGGTCAATTGCTGC 727

Qy      541 CGGCAAACGGGCGTTTCGGGAACTGGGAACGGGGTGGCTTAATTTTCCGCAATCAAACTTGTGC 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      728 CGGCAAACGGGCGTTTCGGGAACTGGGAACGGGGTGGCTTAATTTTCCGCAATCAAACTTGTGC 787

Qy      601 CTTAAACGTGAGAAACAGGACGCAAAAGCTGTTCATTAACGGCATTACCATCGGGGAACTC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      788 CTTAAACGTGAGAAACAGGACGCAAAAGCTGTTCATTAACGGCATTACCATCGGGGAACTC 847

Qy      661 AAGGTGAGCGCGCCGACAAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      848 AAGGTGAGCGCGCCGACAAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 907

Qy      721 GAGATATTTCTCCTGCAACAAGCACAGGGATGCCGGAGCCGGGGGAAGGGGATCACC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      908 GAGATATTTCTCCTGCAACAAGCACAGGGATGCCGGAGCCGGGGGAAGGGGATCACC 967

Qy      781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAAATTTATTTGCTACAA 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      968 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAAATTTATTTGCTACAA 1027

Qy      841 CGCACGCGCCAGGAGGTTGCCCGCAGCCGCGCCACCCCGGTTGATCAAGACACGG 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1028 CGCACGCGCCAGGAGGTTGCCCGCAGCCGCGCCACCCCGGTTGATCAAGACACGG 1087

Qy      901 TTGACGCGCCCATCCACCCGACCAAAAAACAGGCGTATGGGGTGTGACATTACCCACTTCAGTGCTG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1088 TTGACGCGCCCATCCACCCGACCAAAAAACAGGCGTATGGGGTGTGACATTACCCACTTCAGTGCTG 1147

Qy      961 TTTATTGCCGGGACACGATACTAATCTCGGCGGCGCCACTGGAGCTCAAACTGG 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1148 TTTATTGCCGGGACACGATACTAATCTCGGCGGCGCCACTGGAGCTCAAACTGG 1207
```

```
Db   1148  TTTATCGCCGGACACGATACTAATCTGGGCAAATCTCGGCGGGGCGCACTGGAGCTCAACTGG
                 ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1021  ACGCTTCCCGGTCAGCCGGATAACACGCCGCCGCCAGGTGGTGAACTGGGTGTTTGAACGCTGG
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208  ACGCTTCCCGGTCAGCCGGATAACACGCCGCCGCCAGGTGGTGAACTGGGTGTTTGAACGCTGG

Qy   1081  CGTCGGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268  CGTCGGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG

Qy   1141  CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCCCCGGAGGAGGTGAAACTGACC
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328  CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCCCCGGAGGAGGTGAAACTGACC

Qy   1201  CTGGCAGGATGTGAAGAGCGAAAATGCCGAGGGCAATGTGTTCGTTGGGCAGGTTTTACGGCAA
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388  CTGGCAGGATGTGAAGAGCGAAAATGCCGAGGGCAATGTGTTCGTTGGGCAGGTTTTACGGCAA

Qy   1261  ATCGTGAATGAAGCACGGCATACCGGCGGTGCCAGTTTGAGAT
           |||||||||||||||||||||||||||||||||||||| | |
Db   1448  ATCGTGAATGAAGCACGGCATACCCGGCGGTGCCAGTTTGTAAT
```

THIS PAGE LEFT BLANK

us-10-601-319-1.rst

```
GenCore version 5.1.9
Copyright (c) 1993 - 2006  Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on:               June 14, 2006, 10:01:41 ; Search time 6946.38 Seconds
                                                (without alignments)
                                                10650.337 Million cell updates/sec

Title:                US-10-601-319-1
Perfect score:        1322
Sequence:             1 atgaaagcgatcttaatccc..........atcaccaccaccactaa 1323
Scoring table:        IDENTITY_NUC
                      Gapop 10.0 , Gapext 1.0
Searched:             48336798 seqs, 2795665780 residues
Total number of hits satisfying chosen parameters:    96473596

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       EST:*
    1:  gb_est1:*
    2:  gb_est3:*
    3:  gb_est4:*
    4:  gb_est5:*
    5:  gb_est5:*
    6:  gb_htc:*
    7:  gb_est2:*
    8:  gb_est7:*
    9:  gb_est8:*
   10:  gb_est9:*
   11:  gb_gss1:*
   12:  gb_gss2:*
   13:  gb_gss3:*
   14:  gb_gss4:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	614.4	46.5	616	2	BJ074127	BJ074127 BJ074127
c 2	589.8	44.6	853	13	CL662734	CL662734 PRI0142b_
c 3	528.2	40.0	746	4	CA930060	CA930060 SCCCL200_
c 4	521	39.4	529	3	BU713770	BU713770 SJAABUG02
c 5	481.8	36.4	595	8	CF326092	CF326092 JMT1--05-
c 6	476.8	36.1	895	8	CN762997	CN762997 ID0AAA5DC
c 7	438.8	33.2	868	8	CN754382	CN754382 ID0AAA12D
c 8	434.8	32.9	864	8	CN757565	CN757565 ID0AAA1DF
c 9	386.2	29.2	393	7	BE520240	BE520240 M1lB12STM
c 10	383	29.0	383	7	AW036132	AW036132 EST274508
c 11	383	29.0	383	7	AW036134	AW036134 EST274510
c 12	348.6	26.4	531	3	BU713769	BU713769 SJAABUG01
c 13	304.6	23.0	354	1	AA545747	AA545747 HBMSF1B4-
c 14	295.8	22.4	706	2	CN759004	CN759004 ID0AAA24B
c 15	248	18.8	707	2	BJ619443	BJ619443 BJ619443
c 16	216.8	16.4	320	7	BE436403	BE436403 EST407481
c 17	208	15.7	208	2	BM409016	BM409016
c 18	208	15.7	208	2	BM412806	BM412806 EST583133
c 19	198.8	15.0	546	2	BJ040796	BJ040796 BJ040796
```

ALIGNMENTS

c 20	169.4	12.8	7	179	7	BE461872	BE461872 EST413291
c 21	160.8	12.2	5	508	2	BJ030203	BJ030203 BJ030203
c 22	157.6	11.9	5	657	9	BG457132	BG457132 NF10OC005P
c 23	156.4	11.8	8	1171	8	CV672454	CV672454 RET7SJ_01
c 24	143.2	10.8	8	254	7	BE520241	BE520241 M1lB12XTM
c 25	103.4	7.8	8	769	8	CO386885	CO386885 AGENCOURT
c 26	88.8	6.7	8	863	9	CV672453	CV672453 RET7SJ_01
c 27	57.2	4.3	14	772	14	DX010757	DX010757 KBrB001C0
c 28	55.2	4.2	13	821	13	CL665831	CL665831 PRI014d_E
c 29	49.8	4.2	13	842	13	CL665745	CL665745 PRI014c_E
c 30	48.8	3.7	9	470	9	DN201839	DN201839 USDA-FP_I
c 31	43.8	3.3	9	392	9	CX943192	CX943192 LamjGestG
c 32	41.2	3.1	2	733	2	BI152280	BI152280 60297706
c 33	39.8	3.0	2	346	4	BY053637	BY053637 BY053637
c 34	39.8	3.0	4	350	4	BY041909	BY041909 BY041909
c 35	39.8	3.0	4	365	4	BY056036	BY056036 BY056036
c 36	39.8	3.0	4	389	4	CA566872	CA566872 K0406H12-
c 37	39.8	3.0	4	431	1	AI553154	AI553154 vi07b12.y
c 38	39.8	3.0	1	539	4	BX517075	BX517075 BX517075
c 39	39.8	3.0	4	602	1	BB610221	BB610221 BB610221
c 40	39.8	3.0	1	607	1	AA521521	AA521521 vi07b12.r
c 41	39.8	3.0	1	632	4	BY738708	BY738708 BY738708
c 42	39.8	3.0	4	646	4	BY737657	BY737657 BY737657
c 43	39.8	3.0	6	2863	6	AK146141	AK146141 Mus muscu
c 44	39.8	3.0	6	2863	6	AK168364	AK168364 Mus muscu
c 45	39.8	3.0	6	2863	6	AK168364	AK168364 Mus muscu
```

RESULT 1
BJ074127/c

```
LOCUS       BJ074127                616 bp    mRNA    linear   EST 29-SEP-2003
DEFINITION  BJ074127 NIBB Mochii normalized Xenopus tailbud library Xenopus
            laevis cDNA clone XL090J06 5', mRNA sequence.
ACCESSION   BJ074127
VERSION     BJ074127.1  GI:37504316
KEYWORDS    EST.
SOURCE      Xenopus laevis (African clawed frog)
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Xenopus.
REFERENCE   1  (bases 1 to 616)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and
            Kohara,Y.
  TITLE     Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp
            The information of this clone is available through the following
            URL.
            http://xenopus.nibb.ac.jp.
FEATURES             Location/Qualifiers
     source          1..616
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8355"
                     /clone="XL090J06"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 25"
                     /clone_lib="NIBB Mochii normalized Xenopus tailbud
                     library"
ORIGIN
Query Match             46.5%;  Score 614.4; DB 2; Length 616;
Best Local Similarity   99.8%;  Pred. No. 8.3e-180;
```

Matches 615; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy   264 GGTAGCGCCACGGATTGCTGGGCGAAAAAAGGGGCTCCCCCGAGTCTGGTCAGGTCGGATTAT 323
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   616 GGTAGCCGACGGATTGCTGGGCGAAAAAAGGGGCTGCCCCCAGTCTGGTCAGGTCGCGGATTAT 557

Qy   324 TGCTGAGTGCGACGGAGGCTGTACCCGTAAAAACAGGCGGCCCTTCGCCCGCCGCCCGGGCTCGCACC 383
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   556 TGCTGAGTGCGACGGAGGCCTACCCGTAAAAACAGGCGAAGCCTTCGCCCGGGGCTCGGCACC 497

Qy   384 TGACTGTGCAATAAACCGTTACAATACCCAGGCAGATACGTCCAGTCCCGATCCGTTATTTAA 443
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   496 TGACTGTGCAATAAACCGTTACAATACCCAGGCAGATACGTCCCAGTCCCGATCCGTTATTTAA 437

Qy   444 TCCTCTAAAAAACTGGCGTTTGCCAACTGGATAACGCGAAACGTGACTGACGCGGATCCTCAG 503
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   436 TCCTCTAAAAAACTGGCGTTTGCCAACTGGATAACGCGAAACGTGACTGACGCGGATCCTCAG 377

Qy   504 CAGGGCAGGAGGGGTCAATTGCTCGACTTTTACCCGGGCCATCGGGCAAAACGCGTTTCGCGAACT 563
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   376 CAGGGCAGGAGGGGTCAATTGCTCGACTTTTACCCGGGCCATCGGGCAAAACGCGTTTCGCGAACT 317

Qy   564 GGAACGGGTGCTTAATTTTCCCGCAATCAAAACTTGTGTGCCTTAAAACGTGAGAAACAGGACGA 623
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   564 GGAACGGGTGCTTAATTTTCCCGCAATCAAAACTTGTGTGCCTTAAAACGTGAGAAACAGGACGA 623

Qy   316 GGAACGGGTGCTTAATTTTCCCGCAATCACTTGTGCCTTAAAACGTGAGAAACAGGACGA 257

Qy   624 AAGCTGTTCATTAACGCAGAGGCATTACCATCAAAGGTGAGCGCCGACAATGTCTC 683
             |||||||||||||||||||||||||||||||||||||||||||||||
Db   256 AAGCTGTTCATTAACGCAGAGGCATTACCATCGGAACTCAAAGGTGAGCGCCGACAATGTCTC 197

Qy   684 ATTAACCGGTGCGGTAAGCCTTGGCATCAAATGCTGACGGAGATATTTCTTCCTGCACAAGC 743
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   196 ATTAACCGGTGCGGTAAGCCCTGCCATCAAATGCTGACGGAGATATTTCTTCCTGCACAAGC 137

Qy   744 ACAGGGGAATGCCGGAGCCCGGGTGGAAGGATCACCAGTTCACACCAGTGGAACACCTT 803
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   136 ACTGGGGAAATGCCGGAGCCCGGGTGGGGGAAGGAATCACCGATTCACACCAGTGGAACACCTT 77

Qy   804 GCTAAGTTTGCATAACGGGCAATTTTTATTTTGCTACACAACGGACAGCCCAGAGGGTTGCCCCGCAG 863
             |||||||||||||||||||||||||||||||||||||||||||||
Db   76  GCTAAGTTTGCATAACGGGCAATTTTTATTTTGCTACACAACGGACAGCCCAGAGGGTTGCCCCGCAG 17

Qy   864 CCGCGCCACCCCGTTA 879
             ||||||||||||||||
Db   16  CCGCGCCACCCCGTTA 1
```

RESULT 2
LOCUS   CL662734
DEFINITION PRIO142b_B11 - PRIO142b_B21 (853) Mixed stage fosmid library of P.
            pacificus var. California Pristionchus pacificus genomic, genomic
            survey sequence.
ACCESSION   CL662734
VERSION   CL662734.1   GI:50150877
KEYWORDS   GSS.
SOURCE   Pristionchus pacificus
  ORGANISM Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1 (bases 1 to 853)
  AUTHORS   Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
  TITLE   AppaDB: an AcedB database for the nematode satellite organism
            Pristionchus pacificus
  JOURNAL   Nucleic Acids Res. 32 (1), D421-D422 (2004)
   PUBMED   14681447
  COMMENT   Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de

FEATURES   Location/Qualifiers
     source   1..853
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Mixed stage fosmid library of P. pacificus
                     var. California"
                     /note="Vector: pEpifos-5 Fosmid vector"

ORIGIN

Query Match   44.6%; Score 589.8; DB 13; Length 853;
Best Local Similarity 99.7%; Pred. No. 4.3e-172;
Matches 591; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy   708 ATCAAATGCTGACGGAGATATTTCTTCCTGCACAAGCACAGGGGAATGCCGGAGCCCGGGGTG 767
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATCAAATGCTGACGGAGATATTTCTTCCTGCAACAAGCACAGGGGAATGCCGGAGCCCGGGGTG 60

Qy   768 GGGAAGGAATCACCGGATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGGGCAATT 827
             ||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GGGAAGGAATCACCGGATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGGGCAATT 120

Qy   828 TTATTTTGCTACAACGCACGCAGAGGGTTGCCCGCCAGCCGCGCCACCCCGTTATTTATTTGGATTT 887
             |||||||||||||||||||||||||||||||||||||||||||
Db   121 TTATTTTGCTACAACGCACGCCAGAGGGTTGCCCGCCAGCCGCGCCACCCCGTTATTGGATTT 180

Qy   888 GATCATGGCAGCGTTGACGCCCCCATCCACCCGGCAAAAACAGGGGTTATGGGTGAGACATTACC 947
             |||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GATCATGGCAGCGTTGACGCCCCCATCCACCCGAAAAAACAGGGCGTATGGGTGTGACATTACC 240

Qy   948 CACTTCAGTACTGTTTATTGCCGGACGACGATTAATCTGGCAAAATCTTCGGCGGCGACT 1007
             |||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CACTTCAGTACTGTTTATTGCCGGACGACGATTACTAATCTGGCAAAATCTTCGGCGGCGACT 300

Qy  1008 GGAGCTCAAACTGGACGCGTTCCCGGGTCAGCCCGGGGATAACACGGCGCCCAGGTGGTGAACTGGT 1067
             |||||||||||||||||||||||||||||||||||||||||||
Db   301 GGAGCTCAAACTGGACGCGTTCCCGGGTCAGCCCGGGGATAACACGGCGCCCAGGTGGTGAACTGGT 360

Qy  1068 GTTTGAACGGCTGGGCGGTCGGGCTAAGCGATAACAGGCCACTGTGGATTCAGGTTTCGGTGGTCTT 1127
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GTTTGAACGGCTGGGCGGTCGGGCTAAGCGATAACAGGCCACTGTGGATTCAGGTTTTCGGTGGTCTT 420
```

RESULT 3
LOCUS   CA093060
DEFINITION SCCCL2001D10.b CL2 Saccharum officinarum cDNA clone SCCCL2001D10
            3', mRNA sequence.
ACCESSION   CA093060
VERSION   CA093060.1   GI:34946367
KEYWORDS   EST.
SOURCE   Saccharum officinarum
  ORGANISM Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum

LOCUS   CA093060   746 bp   mRNA   linear   EST 23-SEP-2003

```
Qy  1128 CCAGACTTTACAGCAGATGCCGTGATAAAAACGCCGCTGTCATTAAATACGCCGCCCGGAGA 1187
             |||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CCAGACTTTACAGCAGATGCGTGATAAAAACGCCGCTGTCATTAAATACGCCGCCCGGAGA 480

Qy  1188 GGTGAAACTGACCCTGCCAGGATGTGAAGACGGAAATGCGCAGGGCATGTGTTCGTTGGC 1247
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GGTGAAACTGACCCTGCCAGGATGTGAAGACGGAAATGCGCAGGGCATGTGTTCGTTGGC 540

Qy  1248 AGGTTTTACGCAAATCCTTGAATGAAGCACGCGATACCGGGCGTGCAGTTTGAGAT 1300
             |||||||||||||||||||||||||||||||||||||||||||
Db   541 AGGTTTTACGCAAATCGTTGAATGAAGCACGCATACCGGGCGTGCAGTTTGTAAT 593
```

```
REFERENCE    1 (bases 1 to 746)
  AUTHORS     Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE       The libraries that made SUCEST
  JOURNAL     Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
  COMMENT     Contact: Arruda P
              Centro de Biologia Molecular e Engenharia Genetica
              Universidade Estadual de Campinas
              Caixa Postal 6010, 13083-970, Campinas SP, Brazil
              Tel: 55 19 3788 1137
              Fax: 55 19 3788 1089
              Email: parruda@unicamp.br
              Clone distribution information can be found
              through the Brazilian Clone Collection Center (BCCC) at
              http://www.bcccenter.fcav.unesp.br
              Plate: 001  row: D  column: 10
              Seq primer: M13/Forward primer.
FEATURES             Location/Qualifiers
     source          1..746
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCCCL2001D10"
                     /lab_host="XL1Blue MRF'"
                     /clone_lib="CL2"
                     /note="Organ: Pool of sugarcane calli submitted to low
                     (4oC) and high (37 C) temperature stress; Vector:
                     pBluescript; Site_1: EcoRI; Site_2: XhoI; An
                     unidirectional cDNA library generated from [Pool of
                     sugarcane calli submitted to low (4oC) and high (37 C)
                     temperature stress]. cDNA was prepared from polyA+ mRNA
                     using ZAP - cDNA Synthesis Kit (Stratagene). The
                     double-strand cDNAs were fractionated in a sepharose CL-2B
                     40cm-columns and fragments sizing between 0.8 and 1.5 Kb
                     were directionally cloned into the vector. Details of
                     each source of RNA and library construction can be
                     obtained at http://sucest.lad.ic.unicamp.br/public"

ORIGIN
Query Match             40.0%;  Score 528.2;  DB 4;  Length 746;
Best Local Similarity   97.0%;  
Matches  547;  Conservative    0;  Mismatches  16;  Indels   1;  Gaps   1;

Qy        595 TTGTGCCTTAAACGTGAGAAACAGGACGAAAAGCTGTTCATTAACGCAGGCAATTACCATCG 654
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 TTGCGTACTGTGCAGTGAGAAAACAGGACGAAAAGCTGTTCATTAACGCAGGCAATTACCATCG 243

Qy        655 GAACTCAAGGTGAGCGCCGACAAATGTCTCATTAACCGGTGCGTAAGCCTTCGCATCAATG 714
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 GAACTCAAGGTGAGCGCCGACAAATGTCTCATTAACCGGTGCGTAAGCCTTCGCATCAATG 303

Qy        715 CTGACGGAGATATTTCTCCTGCAACAAGCACAGGGAAATGCCCGAGCCCGGGGTGGGGAAGG 774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 CTGACGGAGATATTTCTCCTGCAACAAGCACAGGGAAATGCCCGAGCCCGGGGTGGGGAAGG 363

Qy        775 ATCACCGATTCACACCAGTGGAACACCTTGCTAAGTTGCATAACGGGCAATTTATTTG 834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 ATCACCGATTCACACCAGTGGAACACCTTGCTAAGTTGCATAACGGGCAATTTATTTG 423

Qy        835 CTACAACGCACGCCAGAGGTTGCCCGCAGCCCGCCCACCCCCGTTATTGATTTGATCATG 894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CTACAACGCACCCCAGAGGTTGCCCGCAGCCCGCCCACCCCCGTTATTAGATTTGATCAAG 483

Qy        895 GCAGCGTTGAGGCCCCCATCCACCGCAAAAACAGGCGTAATGGTGTGACATTACCCACTTCA 954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 ACAGCGTTGACCGCCCATCCACGCANAAACAGGCGTAATGGTGTGACATTACCCACTTCA 543

Qy        955 GTACTGTTTAATTGCCGGACACGATACTAATCTGGCAANAATCTCGGCGGCGGCACTGGAGCTC 1014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 GTGCTGTTTAATGCGCGACACGATACTAATCTGGCAANAATCTCGGCGGCGGCACTGGAGCTC 603

Qy       1015 AACTGGACGCTTCCCGGTCAGCGGGATAACACGCCCGGCCAGGTGGTGAACTGGTGTTTGAA 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        604 AACTGGACGCTTCCCGGTCAGCGGGATAACACGCCCGGCCAGGTGGTGAACTGGTGTTTGAA 662
Qy       1075 CGCTCGGCGTCGGCTAAGCGGATAACAGCACAGTGAGATTCAGGTTCGCTGGTTCGTTCCACAGACT 1134
              |||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 CGCTCGGCGTCGNGCTAAGCGGATAACAGCACAGTGAGATTCAGGTTCGCTGGTTCGTTCCACAGACT 722
Qy       1135 TTACAGGCAGATGCCTGATAAAACG 1158
              ||||||||| |||||||||||||||
Db        723 TTACAGCANATGCGTGATAAAACG 746
```

```
RESULT 4
BU713770
LOCUS       BU713770          529 bp    mRNA    linear    EST 23-OCT-2003
DEFINITION  SJAABUG02 Adult SjC 7/94 Schistosoma japonicum cDNA similar to
            pdb|1DKP|A Chain A, Crystal Structure Of Phytate Complex Of
            Escherichia Coli Phytase At Ph 6.6. Phytate Is Bound With Its
            3-Phosphate In The Active Site. Hg2+ Cation Acts As An
            Intermolecular Bridge, mRNA sequence.
ACCESSION   BU713770
VERSION     BU713770.1  GI:28321126
KEYWORDS    EST.
SOURCE      Schistosoma japonicum
  ORGANISM  Schistosoma japonicum
            Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
            Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE   1 (bases 1 to 529)
  AUTHORS   Hu,W., Yan,Q., Shen,D.K., Liu,F., Zhu,Z.D., Song,H.D., Xu,X.R.,
            Wang,Z.J., Rong,Y.P., Zeng,L.C., Wu,J., Zhang,X., Wang,J.J.,
            Xu,X.N., Wang,S.Y., Fu,G., Zhang,X.L., Wang,Z.Q., Brindley,P.J.,
            McManus,D.P., Xue,C.L., Feng,Z., Chen,Z. and Han,Z.G.
  TITLE     Evolutionary and biomedical implications of a Schistosoma japonicum
            complementary DNA resource
  JOURNAL   Nat. Genet. 35 (2), 139-147 (2003)
  PUBMED    12973349
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn.
FEATURES             Location/Qualifiers
     source          1..529
                     /organism="Schistosoma japonicum"
                     /mol_type="mRNA"
                     /strain="Chinese (Anhui) strain"
                     /db_xref="taxon:6182"
                     /sex="Male and female"
                     /tissue_type="Whole body"
                     /dev_stage="Adult worms"
                     /lab_host="Mouse and rabbit"
                     /clone_lib="Adult SjC 7/94"
                     /note="Vector: Lambda ZAP-II XR.; Site 1: EcoR I; Site_2:
                     XhoI I; Several hundred adult Schistosoma japonicum
                     (Anhui, P.R. China, strain), of mixed sex, were perfused
                     from the mesenteries of experimentally infected mice and
                     rabbits at the Queensland Institute of Medical Research,
                     Brisbane, Australia (QIMR), and stored for several months
                     in liquid nitrogen. Subsequently, mRNA was isolated at the
                     QIMR from lysates of these worms by oligo dT
                     chromatography, using a kit from Pharmacia. The mRNA was
                     then shipped to Clontech, Palo Alto, CA, USA, who
                     constructed a cDNA library. First strand synthesis was
                     primed with an oligo-dT-XhoI-primer and synthesized using
                     M-MLV reverse transcriptase. Second strand synthesis was
                     accomplished with RNase H and T4 DNA polymerase. The
                     double stranded cDNA was ligated to EcoRI linkers,
                     digested with EcoRI and XhoI, and ligated into the
                     phagemid vector lambda ZAP II XR. After construction of
                     this directional library by Clontech, it was returned to
                     the QIMR. During analysis of the library at the QIMR, we
                     have found that a small percentage, 2% to 3%, of the
```

COMMENT   clones contain inserts that appear to be highly homologous
                to sequences from salmonid fishes, as determined by
                homology comparisons using BLAST and by Southern
                hybridization analysis to genomic DNA from salmon (Sigma
                Chemical Co., St. Louis, MO) under stringent washing
                conditions. The remainder of the clones appear to contain
                S. japonicum sequences."

ORIGIN
Query Match   39.4%;   Score 521; DB 3; Length 529;
Best Local Similarity 99.1%;   Pred. No. 1.2e-150;
Matches 524; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy   670 GCCGACAAATGTTCTCATTAACCGGTGCGGGTAGGCTGCGCATCAATGCTGAGACCGAGATATTT 729
          |||||||||||| ||||||||||||||||||||||||| |||||||| |||||||| |||||||||
Db   1 GCCGACAAATGTTCATTAACCGGTGCGGGTAGGCCTCGCATCAAATGCTGACCGAGATATTT 60

Qy   730 CTCCTGCAACAAGCACAGGGAATGCCCGAGCCCGGGGTGGGGAAGGATCACCGGATTCACAC 789
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTCCTGCAACAAGCACAGGGAATGCCCGAGCCCGGGGTGGGGAAGGATCACCGGATTCACAC 120

Qy   790 CAGTGGAACAACCTTGCTAAGTTGCAATAACCGCGCAATTTTTATTTGCTACACGCACGCCA 849
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CAGTGGAACAACCTTGCTAAGTTGCATAACCGCGCAATTTATTATTTGCTACACGCACGCCA 180

Qy   850 GAGGGTTGCCCCGGCAGCCCACCCCGTTATTGGATTTGATCATGGCAGCGGTTGACGCGCC 909
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GAGGGTTGCCCCGGCAGCCCACCCCGTTATTGATTTGATCATGGCAGCGGTTGACGCGCC 240

Qy   910 CATCCACCGCAAAAACAGGGCGTATGGTGTGACATTACCACTTCAGTGTCGTTTATTATTGCC 969
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CATCCACCGCAAAAACAGGGCGTATGGTGTGACATTACCCACTTCAGTGTCGTTTATCGCC 300

Qy   970 GGACACGGATACTATAATCTGGCAAATCTCGGCGCACTGGGAGCTCCGACGCTTCCC 1029
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GGACACGGATACTATAATCTGGCAAATCTCGGCGCACTGGGAGCTCCGACGCTTCCC 360

Qy 1030 GGTCAGCGCGGATAACACGGCCCGGCAGTGGGTGGTGGTGAACTGGTGTTTGTTTGAACCGCTCGGCGGTCGGCTA 1089
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGTCAGCGCGGATAACACGGCCCGGCAGTGGGTGGTGGTGAACTGGTGTTTGTTTGAACCGCTCGGCGGTCGGCTA 420

Qy 1090 AGCGGATAACAGCCAGTGGGATCAGGTTTCGCTGCGTCTTCCAGACTTTACAGCAGATGCGGT 1149
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AGCGGATAACAGCCAGTGGGATTCAGGTTTCGCTGCGGTCTTCCAGACTTTACAGCAGATCGGT 480

Qy 1150 GATAAAAACGCGCGCTGTCATTAAATAACCCGCGCCCGGGAGGGTGAGGGTGAAAACTGA 1198
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GATAAAAACGCGCGCTGTCATTAAATAACGCGCGCCCGGGAGGGTGAGGGTGAAAACTGA 529

RESULT 5
LOCUS   CF326092   595 bp   mRNA   linear   EST 18-AUG-2003
DEFINITION   JMT1--05-B09.g1 AtJMT-overexpressing transgenic rice lambda
              cDNA library (JMT1) Oryza sativa (japonica cultivar-group) cDNA
              clone JMT1--05-B09), mRNA sequence.
ACCESSION   CF326092
VERSION   CF326092.1 GI:33800445
KEYWORDS   EST.
SOURCE   Oryza sativa (japonica cultivar-group)
ORGANISM   Oryza sativa (japonica cultivar-group)
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
              clade; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1 (bases 1 to 595)
AUTHORS   Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
              Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE   Large-scale Sequencing Analysis of Rice ESTs
JOURNAL   Unpublished (2003)
COMMENT   Contact: Nahm B.H.
              Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
              of Bioscience and Bioinformatics, Myongji University
              Yongin, Kyeonggi, Korea

FEATURES   Location/Qualifiers
source   1..595
                /organism="Oryza sativa (japonica cultivar-group)"
                /mol_type="mRNA"
                /cultivar="Nackdong"
                /db_xref="taxon:39947"
                /clone="JMT1--05-B09"
                /tissue_type="leaf"
                /dev_stage="14 days after germination"
                /lab_host="E.coli SOLR"
                /note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
                XhoI; cDNA library "AtJMT-overexpressing transgenic rice lambda
                phage cDNA library (JMT1)"; Site 1: EcoRI; Site_2:
                XhoI; cDNA was inserted into lamda Uni-ZAP XR vector at 5'
                end with EcoRI and 3' end with XhoI site. mRNA was
                prepared from Arabidopsis Jasmoinate Carboxyl
                methyltransferase overexpression line."

ORIGIN
Query Match   36.4%;   Score 481; DB 5; Length 595;
Best Local Similarity 98.6%;   Pred. No. 2.1e-138;
Matches 486; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy   808 AGTTTGCATAACGCGCAATTTTATTTGCTACAAGCGCACGCACGCCAGAGGGTTGCCCGGCGGC 867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   595 AGTTTGCATAACGCGCAATTTTATTTGCTACAAGCGCACGCACGCCAGAGGGTTGCCCGGCGGC 536

Qy   868 GCCACCCCGTTATTGGATTTGATCATGGCAGCGGTTGACGCGCCCATCCACCGCAAAAAACAG 927
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   535 GCCACCCCGTTATTGATTTGATCATGGCAGCGGTTGACGCGCCCATCCACCGCAAAAAACAG 476

Qy   928 GGCGTATGGTGTGACATTACCCACTTCAGTACTGTTTATTGCCGGACACGGATACTATAATCTG 987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   475 GGCGTATGGTGTGACATTACCCACTTCAGTGCTGTTTATTGCCGGACACGGATACTATAATCTG 416

Qy   988 GCAAATCTCGGCGCGCGCACTGGGAGCTCAACTGGACAGCTTCCCGGTCAGCGCGGATAACACG 1047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   415 GCAAATCTCGGCGCGCGCACTGGGAGCTCAACTGGACAGCTTCCCGGTCAGCGCGGATAACACG 356

Qy 1048 CCGCCAGTGGGTGGTGGTGAACTGGTGTTTGTTTGAACGGCTCGGCGGTCGGCTAAGCGGATAACAGCCAGTGG 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   355 CCGCCAGTGGGTGGTGAACTGGTGTTTGTTTGAACGGCTCGGCGGTCGGCTAAGCGGATAACAGCCAGTGG 296

Qy 1108 ATTCAGGTTTCGCTGGGTCTTCCAGACTTTACAGCAGATGCGGTGATAAAAACGCGCGCTGTCA 1167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   295 ATTCAGGTTTCGCTGGGTCTTCCAGACTTTACAGCAGATGCGGTGATAAAAACGCGCGCTGTCA 236

Qy 1168 TTAAATAACGCCGCCCGCCGGGAGAGGGTGAAACTGAACCCTGGCCAGGAGATGTGAAGAGGGAAAATGCG 1227
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   235 TTAAATAACGCCCGCCGGGAGAGGGTGAAACTGACCCCTGGCCAGGAGATGTGAAGAGCGAAAATCGG 176

Qy 1228 CAGGGCCATGTGTTCGTTGGGCAGGTTTTACGCAAATCGTAGGTTTTACGCAAATCGTAGATGAAGAAGACACGTAGCCGGCG 1287
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   175 CAGGGCCATGTGTTCGTTGGGCAGGTTTTACGCAAATCGTAGGTTTTACGCAAATCGTAGATGAAGAAGACACGTAGCCGGCG 116

Qy 1288 TGCAGTTTGAGAT 1300
          ||||||||| ||
Db   115 TGCAGTTTGTAAT 103

RESULT 6
LOCUS   CN762997/c   895 bp   mRNA   linear   EST 20-MAY-2004
DEFINITION   ID0AA5DC09RM1 ApMS Acyrthosiphon pisum cDNA clone ID0AA5DC09 5',
              mRNA sequence.
ACCESSION   CN762997
VERSION   CN762997.1 GI:47536920
KEYWORDS   EST.
SOURCE   Acyrthosiphon pisum (pea aphid)
ORGANISM   Acyrthosiphon pisum

Email: bhnahm@eggbio.com, bhnahm@bio.myongji.ac.kr.
Tel: 82 31 330 6193
Fax: 82 31 321 6355

```
REFERENCE    Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
AUTHORS        Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidoidea;
               Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
             1 (bases 1 to 895)
               Hunter,W., Martinez-Torres,D., Rahbe,Y., Sabater-Munoz,B.,
               Stern,D., Tagu,D. and Wincker,P.
TITLE        An expressed sequence tags database for the pea aphid Acyrthosiphon
               pisum
JOURNAL      Unpublished (2004)
COMMENT      Contact: D. Tagu
               INRA Rennes
               UMR BiO3P, BP 35327, F-35653 Le Rheu Cedex France
               Tel: +33.2.23.48.51.65
               Fax: +33.2.23.48.51.50
               Risk of contamination by bacterial sequences from obligatory
               (Buchnera) or facultative endosymbionts. These sequences were
               obtained in the frame of the International Consortium of Aphid
               Genomics in collaboration with Genoscope
PCR Primers
FORWARD: CAGGAAACAGCTATGACC
Plate: 5    row: C    column: 9.
FEATURES             Location/Qualifiers
source          1..895
                /organism="Acyrthosiphon pisum"
                /mol_type="mRNA"
                /cultivar="developmentstage"
                /db_xref="taxon:7029"
                /clone="IDOAAA5DC09"
                /tissue_type="whole insect"
                /dev_stage="nympha and adults (parthenogenetic females)"
                /lab_host="Xl1-Blue"
                /clone_lib="ApMS"
                /note="Vector: pBS-SK minus; Site_1: EcoRI; Site_2: XhoI;
                Sample name: IDOAAA ; Plant growth place: Department of
                Ecology & Evolutionary Biology, Princeton University ;
                Soil conditions: Soil ; Sowing date: 01/06/1999 ;
                Harvesting date: 01/06/1999 ; Stress date: no stress ;
                Description: Aphids inoculated on one-week old Vicia faba
                under non-sterile conditions. All parthenogenetic stages
                and both winged and wingless adults were collected for
                library construction. ; experimental condition: long
                photoperiod (16-hr light/8-hr dark at 18 c)"
ORIGIN

Query Match             36.1%;     Score 476.8;    DB 8;     Length 895;
Best Local Similarity   98.6%;
Matches    481;    Conservative    0;    Mismatches    7;    Indels    0;    Gaps    0;

Qy    813  GCATAACGCGCAATTTTTATTTTGCTACAACGCACGCCCAGAGGGTTGCCCGCCGCGCCAC  872
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    895  GCATAACGCGCAATTTTTATTTTGCTACAACGCACGCCCAGAGGGTTGCCCGCCGCGCCAC  836

Qy    873  CCCGGTTATTGATTGGATTCATGGCAGCCGGTTGACGGCCCATCCACCGCAAAACAGGCGTA  932
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    835  CCCGGTTATTGATTGGATTCATGGCAGCCGGTTGACGGCCCATCCACCGCAAAACAGGCGTA  776

Qy    933  TGGTGTGACATATACCCACTTCAGTACTGTTTATTGCCGGACACGGATACTAATCTGGCAAA  992
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    775  TGGTGTGACATTACCCACTTCAGTCGTCTTTATCGCCGGACACGGATACTAATCTGGCAAA  716

Qy    993  TCTCGGCGGGGCGCACTGAGCTCAAACTGGACCGCTTCCCGGTCAGCCGGATAACAGCCGCC  1052
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    715  TCTCGGCGGGGCGCACTGAGCTCAACTGGACCGCTTCCCGGTCAGCCGGATAACAGCCGCC  656

Qy    1053 AGGTGGTGGAACTGGGTGTTTGGAACGCTGGGGCGTCGGCGTCGGATAACAGCCAGTGGATTCA  1112
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    655  AGGTGGTGAACTGGGTGTTTGGAACGCTGGGGCGTCGGCGGATAACAGCCAGTGGATTCA  596

Qy    1113 GGTTTCGCTGGTCTTCCAGACTTTACAGCAGATGCGTGATAAAAACGCCGCTGTCATTAAA  1172
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    595  GGTTTCGCTGGTCTTCCAGACTTTACAGCAGATGCGTGATAAAAACGCCGCTGTCATTAAA  536

Qy    1173 TACGCCGCCCCGGAGAGTGAAACTGACCCTGGCAGGATGTGAAGAGACGGAAATGCGCAGGG  1232
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    535  TACGCCGCCCGGAGAGGTGAAACTGACCCTGGCAGGATGTGAAAGTGAAGAGACGGAAATGCGCAGGG  476

Qy    1233 CATGTGTTCGTTGGCGAGGTTTTTACGCAAAATCGTGAATGAAGCAGCAGCCATACCGGCCTGCG  1292
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    475  CATGTGTTCGTTGGCGAGGTTTTTACGCAAAATCGTGAATGAAGCACGCATACCGGCCGTGCAG  416

Qy    1293 TTTGAGAT 1300
           |||| ||
Db    415  TTTGTAAT 408
```

```
REFERENCE    Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
AUTHORS        Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidoidea;
               Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
             1 (bases 1 to 868)
               Hunter,W., Martinez-Torres,D., Rahbe,Y., Sabater-Munoz,B.,
               Stern,D., Tagu,D. and Wincker,P.
TITLE        An expressed sequence tags database for the pea aphid Acyrthosiphon
               pisum
JOURNAL      Unpublished (2004)
COMMENT      Contact: D. Tagu
               INRA Rennes
               UMR BiO3P, BP 35327, F-35653 Le Rheu Cedex France
               Tel: +33.2.23.48.51.65
               Fax: +33.2.23.48.51.50
               Risk of contamination by bacterial sequences from obligatory
               (Buchnera) or facultative endosymbionts. These sequences were
               obtained in the frame of the International Consortium of Aphid
               Genomics in collaboration with Genoscope
PCR Primers
FORWARD: CAGGAAACAGCTATGACC
Plate: 12   row: E    column: 1.
FEATURES             Location/Qualifiers
source          1..868
                /organism="Acyrthosiphon pisum"
                /mol_type="mRNA"
                /cultivar="developmentstage"
                /db_xref="taxon:7029"
                /clone="IDOAAA12DE01"
                /tissue_type="whole insect"
                /dev_stage="nympha and adults (parthenogenetic females)"
                /lab_host="Xl1-Blue"
                /clone_lib="ApMS"
                /note="Vector: pBS-SK minus; Site_1: EcoRI; Site_2: XhoI;
                Sample name: IDOAAA ; Plant growth place: Department of
                Ecology & Evolutionary Biology, Princeton University ;
                Soil conditions: Soil ; Sowing date: 01/06/1999 ;
                Harvesting date: 01/06/1999 ; Stress date: no stress ;
                Description: Aphids inoculated on one-week old Vicia faba
                under non-sterile conditions. All parthenogenetic stages
                and both winged and wingless adults were collected for
                library construction. ; experimental condition: long
                photoperiod (16-hr light/8-hr dark at 18 c)"
ORIGIN

Query Match             33.2%;     Score 438.8;    DB 8;     Length 868;
Best Local Similarity   98.3%;
Matches    454;    Conservative    0;    Mismatches    7;    Indels    1;    Gaps    1;

Qy    839  AAGGCACCGCCAGAGGTTGCCCGCAGCCCGCCACCCCGTTATTGGATTTGGATCATGGCAG  898
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    868  AACGCACCGCCAGA-GTTGCCCGCAGCCCGCCACCCCGTTATTAGATTTGATCAAGACAG  810
```

RESULT 8

```
LOCUS        CN757565/c
DEFINITION   ID0AAA1DF09RM1 ApMS Acyrthosiphon pisum cDNA clone ID0AAA1DF09 5',
             mRNA sequence.
ACCESSION    CN757565
VERSION      CN757565.1  GI:47531488
KEYWORDS     EST.
SOURCE       Acyrthosiphon pisum (pea aphid)
ORGANISM     Acyrthosiphon pisum
             Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
             Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
             Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
REFERENCE    1  (bases 1 to 864)
AUTHORS      Hunter,W., Martinez-Torres,D., Rabbe,Y., Sabater-Munoz,B.,
             Stern,D., Tagu,D. and Wincker,P.
TITLE        An expressed sequence tags database for the pea aphid Acyrthosiphon
             pisum
JOURNAL      Unpublished (2004)
COMMENT      Contact: D. Tagu
             INRA Rennes
             UMR BiO3P, BP 35327, F-35653 Le Rheu Cedex France
             Tel: +33.2.23.48.51.65
             Fax: +33.2.23.48.51.50
             Risk of contamination by bacterial sequences from obligatory
             (Buchnera) or facultative endosymbionts. These sequences were
             obtained in the frame of the International Consortium of Aphid
             Genomics in collaboration with Genoscope
             PCR Primers
             FORWARD: CAGGAAACAGCTATGACC
             Plate: 1   row: F   column: 9.
FEATURES     Location/Qualifiers
source       1..864
             /organism="Acyrthosiphon pisum"
             /mol_type="mRNA"
             /cultivar="developmentstage"
             /db_xref="taxon:7029"
             /clone="ID0AAA1DF09"
             /tissue_type="whole insect"
             /dev_stage="nymphs and adults (parthenogenetic females)"
             /lab_host="XL1-Blue"
             /clone_lib="ApMS"
             /note="Vector: pBS-SK minus; Site_1: EcoRI; Site_2: XhoI;
```

```
Sample name: ID0AAA ; Plant growth place: Department of
Ecology & Evolutionary Biology, Princeton University ;
Soil conditions: Soil ; Sowing date: 01/06/1999 ;
Harvesting date: 01/06/1999 / Stress date: no stress ;
Description: Aphids inoculated on one-week old Vicia faba
under non-sterile conditions. All parthenogenetic stages
and both winged and wingless adults were collected for
library construction. ; experimental condition: long
photoperiod (16-hr light/8-hr dark at 18 c)"
```

```
ORIGIN
Query Match      32.9%;  Score 434.8;  DB 8;  Length 864;
Best Local Similarity  98.3%;  Pred. No. 1.1e-123;
Matches  450;  Conservative  0;  Mismatches  7;  Indels  1;  Gaps  1;

Qy   899  CGTTGACGCCCATCCACCCGAAAAACAGGGGTATAGGTGTGAGTGAGACTATTACCCACTTCAGTAC   958
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db   809  CGTTGAGCGCCCATCCACCCGAAAAACAGGGGTATAGGTGTGAGTGAGACTATTACCCACTTCAGTGC  750

Qy   959  TGTTTATTGCCGGACACGATACTAATCTGGCAAATCTTCGGCGCCGGCGGCGGCGGCTCAAACT  1018
          |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db   749  TGTTTATTCGCGGACACGATACTAATCTGGCAAATCTTCGGCGCCGGCGGCGGCGGCTCAAACT   690

Qy  1019  GGACGGCTTCCCGGTCAGCCCGGATAACACGCGCCCGCCCAGTGGTGGGTGAACTGGTGTTTGAACGGCT  1078
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   689  GGACGGCTTCCCGGTCAGCCCGGATAACACGCGCCCGCCCAGTGGTGGGTGAACTGGTGTTTGAACGGCT   630

Qy  1079  GGCGGTCGGCTAAGCCGGATAACAGCCAGTGGATTCAGGTTTCGGTTCTTCCAGAGACTTTAC  1138
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   629  GGCGGTCGGCTAAGCCGGATAACAGCCAGTGGATTCAGGTTTCGGTTCTTCCAGAGACTTTAC   570

Qy  1139  AGCAGATGCGTGATAAAAACGCGCGCTGTCATTAAATACGGCCGCCCCGGAGAGGTGAAACTGA  1198
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   569  AGCAGATGCGTGATAAAAACGCGCGCTGTCATTAAATACGGCCGCCCCGGAGAGGTGAAACTGA   510

Qy  1199  CCCTGGCAGGATGTGAAGAGAGCGGAAAATGCGCAGGGGCATGGTGTTCGTTGGGCAGGTTTTTACGC  1258
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   509  CCCTGGCAGGATGTGAAGAGAGCGGAAAATGCGCAGGGGCATGGTGTTCGTTGGGCAGGTTTTTACGC   450

Qy  1259  AAATCGTGAATGAAGCACACGCATACCGGCGGTGCAGTTTGTTTGAGAT  1300
          |||||||||||||||||||||||||||||||||||||||||||  ||
Db   449  AAATCGTGAATGAAGCACACGCATACCGGCGGTGCAGTTTGTTTGTTAAT   408
```

RESULT 9

```
LOCUS        BB520240/c
DEFINITION   M11B12STM Arabidopsis developing seed Arabidopsis thaliana cDNA
             clone M11B12 5', mRNA sequence.
ACCESSION    BB520240
VERSION      BB520240.1  GI:9778242
KEYWORDS     EST.
SOURCE       Arabidopsis thaliana
ORGANISM     Arabidopsis thaliana (thale cress)
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
             rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE    1  (bases 1 to 393)
AUTHORS      White,J.A., Todd,J., Newman,T., Focks,N., Girke,T., Martinez de
             Ilarduya,O., Jaworski,J.G., Ohlrogge,J. and Benning,C.
TITLE        A new set of Arabidopsis expressed sequence tags from developing
             seeds. The metabolic pathway from carbohydrates to seed oil
JOURNAL      Plant Physiol. 124 (4), 1582-1594 (2000)
PUBMED       11115876
COMMENT      Contact: Benning, C
             Dept. of Biochemistry & Molecular Biology
             Michigan State University
             224 Biochemistry, Michigan State University, East Lansing, MI
             48824, USA
```

```
Qy   843  CACGCCAGAGGTTGCCCGCGGCCGCGGCCCACCCCGTTATTGGATTTGATCATGTGCAGGGTT   902
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   864  CACGCCAGAGGTTGCCCGCGGCCGCGGCCCACCCCGTTATTGGATTTGATCATGTGCAGGGTT   806

Qy   903  GACGCGCCCATCCACCCGAAAAACAGGGGTATAGGTGTTGAGACATTACCCACTTCAGTTACTTGTT   962
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   805  GACGCGCCCATCCACCCGAAAAACAGGGGTATAGGTGTTGAGACATTACCCACTTCAGTTACTTGTT   746

Qy   963  TATTGCCGGACACGATACTAATCTGGCAAATCTTCGGCGGCGCACTGGAGCTCAAACTGGAC  1022
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   745  TATTGCCGGACACGATACTAATCTGGCAAATCTTCGGCGGCGCACTGGAGCTCAAACTGGAC   686

Qy  1023  GCTTCCCGGTCAGCCCGGATAACACGCGCCCGCCCAGTGGTGGTGAACTGGTGTTTGTTGTTGAACGGCTGGCG  1082
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   685  GCTTCCCGGTCAGCCCGGATAACACGCGCCCGCCCAGTGGTGGTGAACTGGTGTTGTTTGAACGGCTGGCG   626

Qy  1083  TCGGCTAAGCCGGATAACAGCCAGTGGATTCAGGTTTCGGTTGTTGTTTCTTCCAGAGACTTTACAGCA  1142
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   625  TCGGCTAAGCCGGATAACAGCCAGTGGATTCAGGTTTCGGTTGTTGTTTCTTCCAGAGACTTTACAGCA   566

Qy  1143  GATGCGTGATAAAAACGCGCGCTGTCATTAAATACGGCCGCCCCGGAGAGGTGAAACTGACCCT  1202
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   565  GATGCGTGATAAAAACGCGCGCTGTCATTAAATACGGCCGCCCCGGAGAGGTGAAACTGACCCT   506

Qy  1203  GGCAGGATGTGAAGAGAGCGGAAAATGCGCAGGGGCATGTGTTCGTTGTTGGGCAGGTTTTTACGCAAAT  1262
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   505  GGCAGGATGTGAAGAGAGCGGAAAATGCGCAGGGGCATGTGTTCGTTGGGCAGGTTTTTACGCAAAT   446

Qy  1263  CGTGAATGAAGCACACGCATACCGGCGGTGCAGTTTGTTTGAGAT  1300
          ||||||||||||||||||||||||||||||||||||||||  ||
Db   445  CGTGAATGAAGCACACGCATACCGGCGGTGCAGTTTGTTTGTTAAT   408
```

RESULT 10

LOCUS   AW036132
DEFINITION EST274508 tomato seed, TAMU Lycopersicon esculentum cDNA clone
            cLEE1B23 similar to periplasmic phosphoanhydride phosphohydrolase
            precursor, putative, mRNA sequence.
ACCESSION   AW036132
VERSION   AW036132.1 GI:5894811
KEYWORDS   EST.
SOURCE   Lycopersicon esculentum
  ORGANISM Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1 (bases 1 to 383)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,
            Holt,I.E., Liang,F., Upton,J., Ronning,C.M., Craven,M.B.,
            Fujii,C.Y., Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C.,

FEATURES   Location/Qualifiers
     source   1..383
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEE1B23"
                     /tissue_type="seeds"
                     /dev_stage="quiescent seed"
                     /lab_host="XL1-Blue MRF'"
                     /clone_lib="tomato seed, TAMU"
                     /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
                     XhoI; cLEE - Tomato Seed EST Library. Directionally cloned
                     cDNAs inserted into pBluescript SK(-) at 5' end with
                     EcoRI and 3' end with XhoI site."

ORIGIN
Query Match   29.0%; Score 383; DB 7; Length 383;
Best Local Similarity   100.0%; Pred. No. 1.3e-107;
Matches 383; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   348 TAAAACAGGCGAAGCCTTCGGCCGCCCCGGCGGCTGGCGCCACCTGACTGTGCAATAACCGGTACATAC 407
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TAAAACAGGCGAAGCCTTCGGCCGCCCCGGCGGCTGGCGCCACCTGACTGTGCAATAACCGGTACATAC 60

Qy   408 CCAGGCAGATAACGTCCCAGTCCCCGATCCGGTTATTTAATCCTCTAAAAAACTGGGCGGTTTGCCA 467
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CCAGGCAGATAACGTCCCAGTCCCCGATCCGGTTATTTAATCCTCTAAAAAACTGGGCGGTTTGCCA 120

Qy   468 ACTGGATAACGCGAACGTGACTGACGCGATCCTCAGCACGCAGGGGCAGGAGGGGTCAATTGCTGA 527
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 ACTGGATAACGCGAACGTGACTGACGCGATCCTCAGCACGCAGGGGCAGGAGGGGTCAATTGCTGA 180

Qy   528 CTTTACCCGGGCATCGGCCAAACGGCGGTTTCGGCGAACTGGAACGGCGTGCTTAATTTTCCGCA 587
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTTTACCCGGGCATCGGCCAAACGGCGGTTTCGGCGAACTGGAACGGCGTGCTTAATTTTCCGCA 240

Qy   588 ATCAAAACTTGTGCCTTAAACGTGAGAAAACAGGACGAAAAGCTGTTCATTAACGGCAGGCATT 647
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ATCAAAACTTGTGCCTTAAACGTGAGAAAACAGGACGAAAAGCTGTTCATTAACGGCAGGCATT 300

Qy   648 ACCATCGGAACTCAAGGTGAGCGCCGACAAATGTCTCATTAACCGGTAAGCCTCGC 707
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ACCATCGGAACTCAAGGTGAGCGCCGACAAATGTCTCATTAACCGGTAAGCCTCGC 360

Qy   708 ATCAATGCTGACGGAGATATTTC 730
          |||||||||||||||||||||||
Db   361 ATCAATGCTGACGGAGATATTTC 383

RESULT 11

LOCUS   AW036134
DEFINITION EST274510 tomato seed, TAMU Lycopersicon esculentum cDNA clone
            cLEE1B23 similar to periplasmic phosphoanhydride phosphohydrolase
            precursor, putative, mRNA sequence.
ACCESSION   AW036134
VERSION   AW036134.1 GI:5894813
KEYWORDS   EST.
SOURCE   Lycopersicon esculentum
  ORGANISM Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.

Tel: 517 355 1609
Fax: 517 353 9334
Email: benning@msu.edu
Michigan State University DNA Sequencing Facility Arabidopsis
Biological Resource Center, The Ohio State University, 309 Botany &
Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX:
614290603 TEL: 614292371.

FEATURES   Location/Qualifiers
     source   1..393
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /ecotype="Columbia"
                     /db_xref="taxon:3702"
                     /clone="M1B12"
                     /tissue_type="seed"
                     /dev_stage="5-13 days after flowering"
                     /lab_host="E.coli"
                     /clone_lib="Arabidopsis developing seed"
                     /note="Organ: Developing seed; Vector: pBluescript SK-;
                     Site_1: EcoRI; Site_2: XhoI"

ORIGIN
Query Match   29.2%; Score 386.2; DB 7; Length 393;
Best Local Similarity   98.7%; Pred. No. 1.3e-108;
Matches 388; Conservative   0; Mismatches   5; Indels   0; Gaps   0;

Qy   319 ATTATTGCTGATGTGACCAGACGTACCCGTAAAACAGGCGAAGCCTTCGCCGCCGGCGCTG 378
          ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   393 ATTATTGCTGATGTCGACGAGCGTACCCGTAANACAGGCGAAGCCTTTGCCCGCCGGGNTG 334

Qy   379 GCACCTGACTGTGCAATAACCGTACATACCCAGGCACAGATACGTCCCAGTCCCCGATCCCGTTA 438
          |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   393 GCACCTGACTGTGCAATAACCGTACATACCCAGGCACAGATACGTCCCAGTCCCCGATCCCGTTA 334

Qy   439 TTTAATCCTCTAAAAAACTGGGCGGTTTGCCAACTGGATAAACGCGGAACGTGACTGACGCGGATC 498
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   333 TTTAATCCTCTAAAAAACTGGGCGGTTTGCCAACTGGATAAACGCGGAACGTGACTGACGCGGATC 274

Qy   499 CTCAGCACGCAGGAGAGGGTGCAATTGCTGACTTTACCCGGGCATCGGCCAAAACGGCGGTTTCGC 558
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   273 CTCAGCACGCAGGAGAGGGTCAATTGCTGACTTTACCCGGGCATCGGCCAAAACGGCGGTTTCGC 214

Qy   559 GAACTGGAACGGGGTGCTTAATTTTCCCGGCAATCAAACTTGTGCCTTAAACGTGAGAAAACAG 618
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   153 GAACTGGAACGGGGTGCTTAATTTTCCCGGCAATCAAACTTGTGCCTTAAACGTGAGAAAACAG 94

Qy   619 GACGAAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTCAAGGTGAGCGCCGACAAT 678
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   93 GACGAAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTCAAGGTGAGCGCCGACAAT 34

Qy   679 GTCTCATTAACCGGTGCGGTAAGCCTCGCCATCA 711
          ||||||||||||| || ||||||||||||||||
Db   33 GTCTCATTAAGGCGCTGCGGTAAGCCCTCGCATCA 1

REFERENCE   1 (bases 1 to 383)
AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,
             Holt,I.E., Liang,F., Upton,J., Ronning,C.M., Craven,M.B.,
             Fujii,C.Y., Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C.,
             Martin,G.B., Tanksley,S.D. and Giovannoni,J.
TITLE   Generation of ESTs from tomato seed tissue
JOURNAL   Unpublished (1999)
COMMENT   Contact: CUGI
             Clemson University Genomics Institute
             Clemson University
             100 Jordan Hall, Clemson, SC 29634, USA
             Email: http://www.genome.clemson.edu/orders/index.html
             5 prime sequence.
FEATURES   Location/Qualifiers
     source   1..383
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEB1823"
                     /tissue_type="seeds"
                     /dev_stage="quiescent seed"
                     /lab_host="XL1-Blue MRF'"
                     /clone_lib="tomato seed, TAMU"
                     /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
                     Xhol; cLEB - Tomato Seed EST Library. Directionally cloned
                     cDNAs inserted into pBluescript SK(-) at 5' end with
                     EcoRI and 3' end with Xhol site."

ORIGIN
Query Match   29.0%; Score 383; DB 7; Length 383;
Best Local Similarity   100.0%; Pred. No. 1.3e-107;
Matches   383; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   348 TAAAACAGGCGAAGCCTTGCGCCGCCGGCGGCGCACCTGACTGTGCAATAACCGTACATAC 407
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   383 TAAAACAGGCGAAGCCTTGCGCCGCCGGCGGCGCACCTGACTGTGCAATAACCGTACATAC 324

Qy   408 CCAGGCAGATACGTCCAGTCCGATCCGTTATTTAATCCTCTAAAAACTGGGCGTTTGCCA 467
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   323 CCAGGCAGATACGTCCAGTCCGATCCGTTATTTAATCCTCTAAAAACTGGGCGTTTGCCA 264

Qy   468 ACTGGATAACCGGAACGTGACTGACCCGGATCCTCAGCAGGGGGTCAATTGCTGA 527
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   263 ACTGGATAACCGGAACGTGACTGACCCGGATCCTCAGCAGGGGGTCAATTGCTGA 204

Qy   528 CTTTACCGGGCCATCGGGAACGGGCGTTTCGGAACGGGGTGCTTAATTTTCCGCA 587
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   203 CTTTACCGGGCCATCGGGAACGGGCGTTTCGGAACGGGGTGCTTAATTTTCCGCA 144

Qy   588 ATCAAAACTTGTGCCTTAAACGTGAGAAACAGAGACGAAAGCTGTTCATTAACCGCAGGGCATT 647
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   143 ATCAAAACTTGTGCCTTAAACGTGAGAAACAGAGACGAAAGCTGTTCATTAACCGCAGGGCATT 84

Qy   648 ACCATCGGAACTCAAGGTGACGCGCCGCCACAAATGTCTCATTAACCGGTCGCGGTAAGCCTCGC 707
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   83 ACCATCGGAACTCAAGGTGACGCGCCGCCACAAATGTCTCATTAACCGGTCGCGGTAAGCCTCGC 24

Qy   708 ATCAATGCTGACGGAGATATTTC 730
          |||||||||||||||||||||||
Db   23 ATCAATGCTGACGGAGATATTTC 1

RESULT 12
LOCUS   BU713769   531 bp   mRNA   linear   EST 23-OCT-2003
DEFINITION SJAABUG01 Adult SjC 7/94 Schistosoma japonicum cDNA similar to
            sp|P07102|PPA_ECOLI PERIPLASMIC APPA PROTEIN PRECURSOR (INCLUDES:
            PHOSPHOANHYDRIDE PHOSPHOHYDROLASE (PH 2.5 ACID PHOSPHATASE) (AP);
            6-PHYTASE], mRNA sequence.
ACCESSION   BU713769
VERSION   BU713769.1 GI:28321125
KEYWORDS   EST.
SOURCE   Schistosoma japonicum
  ORGANISM Schistosoma japonicum
            Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
            Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE   1 (bases 1 to 531)
  AUTHORS   Hu,W., Yan,Q., Shen,D.K., Liu,F., Zhu,Z.D., Song,H.D., Xu,X.R.,
            Wang,Z.J., Rong,Y.P., Zeng,L.C., Wu,J., Zhang,X., Wang,J.J.,
            Xu,X.N., Wang,S.Y., Fu,G., Zhang,X.L., Feng,Z., Chen,Z. and Han,Z.G.
            McManus,D.P., Xue,C.L., Peng,Z.,
  TITLE   Evolutionary and biomedical implications of a Schistosoma japonicum
            complementary DNA resource
  JOURNAL   Nat. Genet. 35 (2), 139-147 (2003)
  PUBMED   12973349
COMMENT   Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-5081922
            Email: hanzg@chgc.sh.cn.
FEATURES   Location/Qualifiers
     source   1..531
                     /organism="Schistosoma japonicum"
                     /mol_type="mRNA"
                     /strain="chinese (Anhui) strain"
                     /db_xref="taxon:6182"
                     /sex="Male and female"
                     /tissue_type="Whole body"
                     /dev_stage="Adult worms"
                     /lab_host="Mouse and rabbit"
                     /clone_lib="Adult SjC 7/94"
                     /note="Vector: Lambda ZAP-II XR.; Site 1: EcoR I; Site 2:
                     Xho I; Several hundred adult Schistosoma japonicum
                     (Anhui, P.R. China, strain), of mixed sex, were perfused
                     from the mesenteries of experimentally infected mice and
                     rabbits at the Queensland Institute of Medical Research,
                     Brisbane, Australia (QIMR), and stored for several months
                     in liquid nitrogen. Subsequently, mRNA was isolated at the
                     QIMR from lysates of these worms by oligo dT
                     chromatography, using a kit from Pharmacia. The mRNA was
                     then shipped to Clontech, Palo Alto, CA, USA, who
                     constructed a cDNA library. First strand synthesis was
                     primed with an oligo-dT-XhoI-primer and synthesized using
                     M-MLV reverse transcriptase. Second strand synthesis was
                     accomplished with RNase H and T4 DNA polymerase. The
                     double stranded cDNA was ligated to EcoRI linkers,
                     digested with EcoRI and XhoI, and ligated into the
                     phagemid vector lambda ZAP II XR. After construction of
                     this directional library by Clontech, it was returned to
                     the QIMR. During analysis of the library at the QIMR, we
                     have found that a small percentage, 2% to 3%, of the
                     clones contain inserts that appear to be highly homologous
                     to sequences from salmonid fishes, as determined by
                     homology comparisons using BLAST and by Southern
                     hybridization analysis to genomic DNA from salmon (Sigma
                     Chemical Co., St. Louis, MO) under stringent washing
                     conditions. The remainder of the clones appear to contain
                     S. japonicum sequences."

ORIGIN
Query Match   26.4%; Score 348.6; DB 3; Length 531;
Best Local Similarity   88.0%; Pred. No. 7.9e-97;
Matches   425; Conservative   0; Mismatches   54; Indels   4; Gaps   4;

Qy   670 GCCGACAATGTCTCATTAACCGGTCGCGGTAAGCCTCGGCATCAATGCTGACGGAGATATTT 729
          ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   1 GCCGACAATGTCTCATTAACCGGTCGCGGTAAGCCTCGCATCAATGCTGACGGAGATGGGG 60

Qy   730 CTCCTGCAACAAGCACAGGGAATGCCGGAGACCCGGGGTGGGGAGGATCACCGATTCACAC 789
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   61 CTCCTGCAACAAGCACAGGGAATGCCGGAGCCCGGGGTGGGGAGGATCACCGATTCACAC 120

Qy   790 CAGTGGGAACACCTTGCTRAGTTTGCATRAACGGGCCAATTTATTTATTGCTACAACACCACGCGCCA 849

RESULT 13

```
AA545747/c
```

LOCUS   AA545747   354 bp   mRNA   linear   EST 12-MAY-1999
DEFINITION HBMSP1B4-REV Human Bone Marrow Stromal Fibroblast Homo sapiens cDNA
            clone HBMSP1B4 5', mRNA sequence.
ACCESSION   AA545747
VERSION   AA545747.1 GI:2307026
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1 (bases 1 to 354)
AUTHORS   Jia,L., Robey,P., Young,M., Schuler,G., Powell,J., Yang,L.,
            Lennon,G., Hillier,L., Allen,M., Bowles,L., Geisel,S., Kucaba,T.,
            Marra,M., Martin,J., Steptoe,M., Tan,F., Thelsing,B., Bowers,Y.,
            Wylie,T., Waterston,R., Wilson,R. and Francomano,C.
TITLE   WashU-MGB/NHGRI EST Project
JOURNAL   Unpublished (1997)
COMMENT   Contact: Libin Jia
            National Human Genome Research Institute
            Medical Genetics Branch
            10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
            Tel: 301-402-4877
            Fax: 301-496-7157
            Email: libin@helix.nih.gov
            Seq primer: M13 Reverse.

FEATURES   Location/Qualifiers
     source   1..354
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HBMSP1B4"
                     /sex="Male and Female"
                     /tissue_type="bone marrow"
                     /cell_type="stromal fibroblast"
                     /dev_stage="mixed"
                     /lab_host="XL1-Blue"
                     /clone_lib="Human Bone Marrow Stromal Fibroblast"
                     /notes="Vector: pBluescript; Site_1: EcoRI; Site_2: XhoI"

ORIGIN
Query Match   23.0%; Score 304.6; DB 1; Length 354;
Best Local Similarity 94.9%; Pred. No. 3.9e-83;

```
Qy        121 CAGTGGGACACACCTTGCTAAGTTTGCATAACGCGCAAATTTATTTGCTACAACGCACGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 GAGGTTGCCGCAGCCGGCGCCCACCCGGTTATTGGATTTGATCANGGAGCGTTGACGCGCC 909

Qy        181 GAGGTTGCCGCAGCCGCGCCCACCCGGTTATTAGATTTGATCAAGACAGCGTTGACGCGCCC 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 CATCCACCGGCAAAAACACGGCGTATNGGGTGTGAGCANTTACCCACTTCAGT-ACTGTTTA 968

Qy        241 CGTTTATCATTATGATTTGCATGTGACAGAGACANTTCCCTCTTGAGTGACGAAAAAAGG 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 CATCCACCGGCAAAAACACGGCGTATNGGGTGTGAGCANTTACCCACTTCAGT-ACTGTTTATTGC 968

Qy        301 CGGCCACCAGCCTGAGTAGATCTGGGAGCTCAACTGACGCCTGTC 360
              |||||||||||||||||||||||||||||||||||||||||||||
Db        969 CGGACACGACTAATCTGGCAAATCTCGGCGCGGGCCACTGGAGACGCGCTTCC 1028

Qy        361 CAGGCAGCCGGATAACACGCCCCGGGAGGACGGGAACTGGATGNTTGNAACCCCTGGCGCGTCGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAGGCAGCCGGATAACACGCCCCGGGAGGACGGGAACTGGATGNTTGNAACCCCTGGCGCGTCGG 420

Qy       1029 CCGTTCAGGCGGGATAACACGCCCGCGGGCGGCCACTGGGAGCTCAACTGG-TGTTTGNAACGGCT-GGCGGTGGG 1086
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1029 CCGTTCAGGCGGGATAACACGCCCGCGGGCGGCCACTGGGAGCTCAACTGG-TGTTTGNAACGGCT-GGCGGTGGG 1086

Qy       1087 CTAAGCGGATAACAAGCCCACTGGGATTCAGGTTTCGGCGTCTTCCAGAACTTTACANCAGATG 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTAAGCGGATAACAAGCCCACTGGGATTCAGGTTTCGGCGTCTTCCAGAAC-TTACANCAGATG 479

Qy       1147 CGT 1149
              |||
Db        480 CGT 482
```

RESULT 14

```
CN759004/c
```

LOCUS   CN759004   706 bp   mRNA   linear   EST 20-MAY-2004
DEFINITION ID0AAA24BC004RM1 ApMS Acyrthosiphon pisum cDNA clone ID0AAA24BC004
            5', mRNA sequence.
ACCESSION   CN759004
VERSION   CN759004.1 GI:47532927
KEYWORDS   EST.
SOURCE   Acyrthosiphon pisum (pea aphid)
ORGANISM   Acyrthosiphon pisum
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
            Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
REFERENCE   1 (bases 1 to 706)
AUTHORS   Hunter,W., Martinez-Torres,D., Rahbe,Y., Sabater-Munoz,B.,
            Stern,D., Tagu,D. and Wincker,P.
TITLE   An expressed sequence tags database for the pea aphid Acyrthosiphon
            pisum
JOURNAL   Unpublished (2004)
COMMENT   Contact: D. Tagu
            INRA Rennes
            UMR BIO3P, BP 35327, P-35653 Le Rheu Cedex France
            Tel: +33.2.23.48.51.65
            Fax: +33.2.23.48.51.50
            Risk of contamination by bacterial sequences from obligatory
            (Buchnera) or facultative endosymbionts. These sequences were
            obtained in the frame of the International Consortium of Aphid
            Genomics in collaboration with Genoscope
            PCR PRimers
            FORWARD: CAGGAAACAGCTATGACC
            Plate: 24 row: C column: 4.

FEATURES   Location/Qualifiers
     source   1..706
                     /organism="Acyrthosiphon pisum"
                     /mol_type="mRNA"
                     /cultivars="developmentstage"
                     /db_xref="taxon:7029"
                     /clone="ID0AAA24BC004"
                     /tissue_type="whole insect"
                     /dev_stage="nymphs and adults (parthenogenetic females)"
                     /lab_host="XL1-Blue"
                     /clone_lib="ApMS"
                     /notes="Vector: pBS-SK minus; Site_1: EcoRI; Site_2: XhoI;
                     Sample name: ID0AAA ; Plant growth place: Department of
                     Ecology & Evolutionary Biology, Princeton University ;
                     Soil conditions: Soil ; Sowing date: 01/06/1999 ;

Matches 335; Conservative   0; Mismatches 16; Indels   2; Gaps   2;

Qy   497 TCCTCAGCGACGGGCACGAGAGGGTCAATTTGCTGACTTTACCGGGCATCGGGCAAACGGGCGGGGTTTC 556
              |||||||||||||||| |||||||||||||||||||||||||||| |||||||||||||||| ||| ||||
Db   351 TCTTCAGCCAGGGCAGGA-GGTCAATTGCTGATTTACCGGGGCATCGGGCAAACGGGGGGTTTC 293

Qy   557 GCGAACTGGAACGGGTCGCTTAAATTTTCCCGAATCAAAACTTGTTGCCTTAAACGTGAGAAAC 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   292 GCGAACTGGAACGGGGTG-TTAATTTTCCCGAATCAAAACTTGTTGCCTTAAACGTGAGAAAC 234

Qy   617 AGGACGAAAGCTGTTCATTAACGCAGGACGCATTACCATCGGAACTCAAGGTGAGGCGCGACA 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   233 AGGACGAAAGCTGTTCATTAACGCAGGACGCATTACCATCGGAACTCAAGGTGAGGCGCGACA 174

Qy   677 ATGTCTCATTAACCCGTGCGGGTAAGCCCTGCATCAAATGCTGACGGAGATATTTCTCCTGC 736
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   173 ATGTNTCATTAACCCGGTGCGGGTAAGCCCTGCATCAAATGCTGACGGAGATATTTNTCCTGC 114

Qy   737 AACAAGCACACAGGGAAATGCCCGGAGCGCGGGGGGTGGGGGAAGGATCACCGGATTCACCAGCAGTGGA 796
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   113 AACAAGCACAGGGAAATGCCCGGAGCGCGGGGGTGGGGGAAGGATCACCGGATTCACCACCAGTGGA 54

Qy   797 ACACCTTGCTAAGTTTGCATAACGCGCAAATTTATTTGCTACAACGCACGCCA 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   53 ACACCTTGCTAAGTTTGCATAACGCGCAAATTTATTTGCTACAACGCACGCCA 1
```

```
ORIGIN

Harvesting date: 01/06/1999 ; Stress date: no stress ;
Description: Aphids inoculated on one-week old Vicia faba
under non-sterile conditions. All parthenogenetic stages
and both winged and wingless adults were collected for
library construction. ; experimental condition: long
photoperiod (16-hr light/8-hr dark at 18 c)"

Query Match     22.4%; Score 295.8; DB 8; Length 706;
Best Local Similarity   99.3%; Pred. No. 2.6e-80;
Matches 297; Conservative  0; Mismatches  2; Indels  0; Gaps  0;

Qy 1002 CGCACTGGAGCTCAACTGGAGCGGCTTCCCGGTCAGCCGGATAACACCACCCCAGGGTGGTGA 1061
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  706 CGCACTGGAGCTCAACTGGAGCGGCTTCCCGGTCAGCCGGATAACACCACCCCAGGGTGGTGA  647

Qy 1062 ACTGGGTGTTTGAACGCGTCGGCGGTCCGGCCTAAGCCGCCAGTGGATTCAGGTTTTCGGCT 1121
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  646 ACTGGGTGTTTGAACGCGTCGGCGGTCGGGCCTAAGCCGCCAGTGGATTCAGGTTTTCGGCT  587

Qy 1122 GGTCTTCCAGACTTTACAGCAGATGCGGTGATAAAACGCCGCTGTCATTAAATACGCCGCC 1181
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  586 GGTCTTCCAGACTTTACAGCGAGATGCCGTGATAAAACCCCGCTGTCATTAAATACGCCGCC  527

Qy 1182 CGGAGAGGTGAAACTGACCCCTGGGCAGGATGCCGAAATGCCGCAGGGCCATGTGTTC 1241
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  526 CGGAGAGGTGAAACTGACCCCTGGGCAGGATGCCGAAATGCCGCAGGGCCATGTGTTC  467

Qy 1242 GTTGGGCAGGTTTTACGGAAATCGTGAATGAAGCACCGCATACCGGCGCTGCCAGTTTGTGTTGTAAT 1300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  466 GTTGGGCAGGTTTTACGGAAATCGTGAATGAAGCACCGCATACCGGCGCTGCCAGTTTGTGTTGTAAT  408
```

```
RESULT 15
LOCUS       BJ619443
BJ619443
            BJ619443 NIBB Mochii normalized Xenopus laevis cDNA clone Xl189d24 5', mRNA sequence.
LOCUS       BJ619443      707 bp    mRNA    linear    EST    01-OCT-2003
DEFINITION  Xenopus laevis NIBB Mochii normalized Xenopus early gastrula library
ACCESSION   BJ619443
VERSION     BJ619443.1  GI:37258203
KEYWORDS    EST.
SOURCE      Xenopus laevis (African clawed frog)
ORGANISM    Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Xenopus.
REFERENCE   1  (bases 1 to 707)
AUTHORS     Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and
            Kohara,Y.
TITLE       Expressed genes in X. laevis embryo
JOURNAL     Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            National Institute of Genetics
            Center For Genetic Resource Information
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp
            The information of this clone is available through the following
            URL.
            http://xenopus.nibb.ac.jp.
FEATURES             Location/Qualifiers
     source          1..707
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8355"
                     /clone="Xl189d24"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 10.5"
                     /clone_lib="NIBB Mochii normalized Xenopus early gastrula
                     library"
ORIGIN
```

```
Query Match     18.8%; Score 248; DB 2; Length 707;
Best Local Similarity   98.0%; Pred. No. 2.2e-65;
Matches 251; Conservative  0; Mismatches  5; Indels  0; Gaps  0;

Qy  750 AATGCCCGGAGCCGGGGGTGGGGAAGGATCACCGATTCACACACCAGTGGAACACCTTGCTAAG  809
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 AATGCCCGGAGCCGGGGGTGGGGAAGGATCACCGATTCACACACCAGTGGAACACCTTGCTAAG   60

Qy  810 TTTGCATAAACGGCGCAATTTATTTGCTACAACGCACGCCCAGAGGGTTGCCCGCCAGGCCGCGCGC  869
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTTGCATAAACGGCGCAATTTATTTGCTACAACGCACGCCCAGAGGGTTGCCCGCCAGGCCGCGCGC  120

Qy  870 CACCCCGTTATTGGATTTGATCATGGCGAGCGGCGGTTGACGCCCCATCCACCGGCAAAAAACAGGC  929
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CACCCCGTTATTGGATTTGATCATGGCGAGCGGCGGTTGACGCCCCATCCACCGGCAAAAAACAGGC  180

Qy  930 GTATGGGTGTGACATTACCCACACTTCAGTACTGTTTATTGCCCGGACACGGATACTAATCTGGC  989
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 GTATGGGTGTGACATTACCCACACTTCAGTACTGTTTATTGCCCGGACACGGATACTAATCTGGC  240

Qy  990 AAAATCTCGGCGCGCGCA 1005
        ||||||||||||||||
Db  241 AAAATCTCGGCGCGCGCA  256

Search completed: June 14, 2006, 14:21:54
Job time : 6948.38 secs
```

OM nucleic - nucleic search, using sw model

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

Run on: June 14, 2006, 10:31:46 ; Search time 40.2281 Seconds
                 (without alignments)
                 7741.348 Million cell updates/sec

Title: US-10-601-319-1
Perfect score: 1322
Sequence: 1 atgaaagcggatcttaatccc..............atcaccaccaccactaa 1323

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0
Searched: 290227 seqs, 117694381 residues

Total number of hits satisfying chosen parameters: 580454

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published Applications NA New:*
                 1: /EMC_Celerra_SID53/pcdata/2/pubpna/US09_NEW_PUB.seq:*
                 2: /EMC_Celerra_SID53/pcdata/2/pubpna/US06_NEW_PUB.seq:*
                 3: /EMC_Celerra_SID53/pcdata/2/pubpna/US07_NEW_PUB.seq:*
                 4: /EMC_Celerra_SID53/pcdata/2/pubpna/US08_NEW_PUB.seq:*
                 5: /EMC_Celerra_SID53/pcdata/2/pubpna/PCT_NEW_PUB.seq:*
                 6: /EMC_Celerra_SID53/pcdata/2/pubpna/US10_NEW_PUB.seq:*
                 7: /EMC_Celerra_SID53/pcdata/2/pubpna/US11_NEW_PUB.seq:*
                 8: /EMC_Celerra_SID53/pcdata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	38	2.9	1309	6	US-10-449-902-20908	Sequence 20908, A
2	36.2	2.9	2879	7	US-11-293-697-2180	Sequence 2180, Ap
3	35	2.6	979	7	US-11-154-977-42	Sequence 42, Ap
4	33.4	2.5	2271	7	US-10-953-349-34412	Sequence 34412, A
5	33.2	2.5	1257	7	US-10-953-349-11004	Sequence 11004, A
6	32.2	2.4	3431	7	US-11-217-529-79920	Sequence 79920, A
7	32	2.4	1396	6	US-11-293-654-37	Sequence 37, Appl
8	31.6	2.4	1489	6	US-10-449-902-8987	Sequence 8987, Ap
9	31.6	2.4	1489	6	US-10-449-902-6188	Sequence 6188, Ap
10	31.6	2.4	1600	6	US-10-449-902-17204	Sequence 17204, A
11	31.4	2.4	3129	6	US-10-449-902-13863	Sequence 13863, A
12	31.2	2.4	2144	6	US-10-449-902-12516	Sequence 12516, A
13	31	2.3	1344	6	US-10-449-902-10760	Sequence 10760, A
14	31	2.3	2791	7	US-11-293-697-707	Sequence 707, App
15	30.8	2.3	1393	6	US-10-953-349-27515	Sequence 27515, A
16	30.8	2.3	1461	6	US-10-449-902-11476	Sequence 11476, A
17	30.8	2.3	3553	6	US-10-449-902-15013	Sequence 15013, A
18	30.2	2.3	422	6	US-10-449-902-965	Sequence 965, App
19	30.2	2.3	2036	6	US-10-449-902-22962	Sequence 22962, A
20	30	2.3	3452	7	US-11-289-10-50	Sequence 50, Appl
21	29.8	2.3	1335	6	US-11-289-10-12155	Sequence 12155, A
22	29.8	2.3	1335	6	US-10-953-349-13409	Sequence 13409, A
23	29.8	2.3	1725	6	US-11-217-522-80383	Sequence 80383, A
24	29.6	2.2	2520	7	US-10-449-902-12996	Sequence 12996, A
25	29.4	2.2	15304	6	US-10-505-928-828	Sequence 828, App
			364	7	US-11-154-977-74	Sequence 74, Appl
26	29.4	2.2	702	6	US-10-449-902-5189	Sequence 5189, Ap
27	29.4	2.2	718	7	US-11-154-977-68	Sequence 68, Appl
28	29.4	2.2	807	7	US-11-154-977-42	Sequence 42, Appl
29	29.4	2.2	956	6	US-11-449-902-11182	Sequence 11182, A
30	29.4	2.2	1161	7	US-11-154-977-36	Sequence 36, Appl
31	29.4	2.2	2071	6	US-10-449-902-24250	Sequence 24250, A
32	29.4	2.2	2754	7	US-11-217-529-5697	Sequence 5697, Ap
33	29.2	2.2	772	6	US-10-449-902-18030	Sequence 18030, A
34	29	2.2	1627	6	US-10-449-902-6979	Sequence 6979, Ap
35	29	2.2	3168	6	US-10-449-902-12853	Sequence 12853, A
36	28.8	2.2	3858	6	US-10-449-902-25001	Sequence 25001, A
37	28.8	2.2	2349	6	US-10-449-902-24808	Sequence 24808, A
38	28.8	2.2	3457	7	US-11-293-697-1224	Sequence 1224, Ap
39	28.8	2.2	11544	6	US-10-449-902-741	Sequence 741, App
40	28.6	2.2	1005	6	US-11-246-405-12	Sequence 12, Appl
41	28.6	2.2	1203	7	US-10-449-902-76896	Sequence 76896, A
42	28.6	2.2	1312	6	US-10-449-902-8467	Sequence 8467, Ap
43	28.6	2.2	1581	6	US-10-449-902-25550	Sequence 25550, A
44	28.6	2.2	1808	6	US-10-449-902-25565	Sequence 25565, A
45	28.4	2.1	1445	6	US-10-449-902-21495	Sequence 21495, A

ALIGNMENTS

RESULT 1
US-10-449-902-20908/c
; Sequence 20908, Application US/10449902
; Publication No. US20060123505A1
GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20908
; LENGTH: 1309
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK071209
; DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-20908

Query Match   2.9%; Score 38; DB 6; Length 1309;
Best Local Similarity   51.5%;
Matches   86; Conservative   0; Mismatches   81; Indels   0; Gaps   0;

Qy   212 CGCGNGGNGTGGTGAGGTGAGCTAATCGCCTATCTCGGACATTACCAACGCCACGCGTCTGGTAGCCG 271
          ||| |||| ||||| | ||| |||| ||| ||| | ||||| | ||| |||| || |||| |||
Db   711 CGCACGCAGGTGGCGCGCGGTTCGCCGACGACGTAGACCTTGAAGACACCCCGCGCCGTGGGTTG 652

Qy   272 ACGGATTGCTTGGCGAAAAAAGGGGTGCCCGCCAGTCTGGTCGCGGTCGCGGATTATTGCTGATG 331
          ||| ||| |||| | | | ||||||| || |||| |||| ||| ||| || || || ||||
Db   651 ACGAACTCCTTGGAGGACGACGACGCGCGCGCCGGAGGACGCGCCGTTCGGCGCGGGCGGGTAG 592

Qy   332 TCGACGAGGGTACCCGTAAAACACGGCGAAGCCTTCGCCCGCCGCGGGCTG 378
          ||| || | ||| || | || | || || || ||| || |||| ||| |
Db   591 ACGAGGCGCGATGTCGTGCGACTCGGCGTCGGCGTCGGCGTCGACGGCGACGCGG 545

RESULT 2
US-11-293-697-2180/c
; Sequence 2180, Application US/11293697

```
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: Patentin Ver. 2.1
;
; SEQ ID NO 2180
; LENGTH: 2879
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-2180

Query Match      2.7%; Score 36.2; DB 7; Length 2879;
Best Local Similarity  51.2%; Pred. No. 0.035;
Matches  83; Conservative  0; Mismatches 79; Indels  0; Gaps  0;

Qy  134 CCAAGGCCACCGAACTGATGACGATGCGAGGAGTCACCCCAGAGCGAACGGCCAACCTGGCCGGGTAA 193
            || ||||| || ||| ||| |  | |  | ||||| | ||||   ||  |  |  | | |||
Db  388 CCGCGGCCTTGCTATTGAGGAAGGAGAGTGTCCCCAAAGGCGTCTTCCGGCCGCCAACGT 329

Qy  194 AACTGGGTTGGCTGACACCGGCGNGGTGGTGAGCTAATCGCCTATCTCGGACATTACCAAC 253
            || || || |  |||||| |   ||| | | |||  |  | |  |  |||  | |  ||
Db  328 GCATGTGTTGGGGGGGAAGTCGCCCCAGCGCGGGGGGCCTGATCATCTTCGGCGGCCGTGAGGTCCGCTC 269

Qy  254 GCCAGCGCGTCTGGTTAGCCGACGGATTGCTGGGCGAAAAAGGGCT 295
            || | || || |||| |||| || || ||||   |  |||| |
Db  268 GGGAACGGCGCTTGGGAGTGCACCGAGCTGGGACACCAGTTGCT 227
```

```
RESULT 3
US-10-953-349-34412/c
Sequence 34412, Application US/10953349
Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: Patentin version 3.3
;
; SEQ ID NO 3412
; LENGTH: 979
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34412

Query Match      48.5%; Score 35; DB 6; Length 979;
Best Local Similarity  48.5%; Pred. No. 0.048;
Matches  95; Conservative  0; Mismatches 101; Indels  0; Gaps  0;

Qy  163 ACCCCAGACGCATGGCCAACCTGGCCGGGTAAAACTGGGTTGGCCTGACACCGCGNGGCTGGGT 222
            || ||| || || || || ||| ||||  |  |  | || |||  |||||   ||  ||
Db  709 ACACCCGCGGAAGGACCGGGACCACCGGACCCTGGCTCACTGGTGAGTGGT 650

Qy  223 GAGCTAATCGCCTATCTCGGACATTACCAACGCGCCAGCCGTCTGTAGCCGACGGATTGCTG 282
            |||| || | | |  | ||| | | |  ||| || || ||| ||| |||| ||  | | |
Db  649 GAGCCGGATCGACGAGACGGCCTTCGCCCGGAGCGCGGGGCCACGCGGCGGGCGT 590

Qy  283 GCGAAAAAGGGCGTTGCCCGCAGTCTCGGTCGGTCAGGTCGCGGATTATTGCTGJANGTCGACGAGCGT 342
            ||  |  |||| | ||  | ||| | |||  | |  |||  | ||| |  |  | | | |  ||| ||
Db  589 GCGGGGCGGAGGCAGTCGCGGGACCGGTGCACGACCGGCGTAGAGGGCTTGAGGGAGGCCGTATT 530

Qy  343 ACCGGTAAAACAGGCG 358
            | ||||  ||| ||||
Db  529 GTCGGTGAGAGAGGCG 514
```

```
RESULT 4
US-10-953-349-11004/c
Sequence 11004, Application US/10953349
Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: Patentin version 3.3
;
; SEQ ID NO 11004
; LENGTH: 2271
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-11004

Query Match      2.5%; Score 33.4; DB 6; Length 2271;
Best Local Similarity  60.4%; Pred. No. 0.3;
Matches  55; Conservative  0; Mismatches 36; Indels  0; Gaps  0;

Qy  1212 TGAAGAGCGGAAATGCGCAGGGGCATGCTGTCGTTCGTTGGCGAAATCGTGAATGA 1271
             |||| |  || |||  |  | | ||| | | |  ||| | || |  || || |||
Db  1734 TGAACCGGACAGTTCGGGTGATAATGTACCTGAGAGAACGGGTTTCTGCAGATCTTGAAGGA 1675

Qy  1272 AGCACGCATACCGGCGTGCAGTTTGAGATCT 1302
             |  | | |  |  || ||| ||| ||  | |
Db  1674 CCTGCGGAAAACTAGTTTGCTCTTTGATCTTT 1644
```

```
RESULT 5
US-11-217-529-79920/c
Sequence 79920, Application US/11217529
Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patentin version 3.3
;
; SEQ ID NO 79920
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-79920

Query Match      2.5%; Score 33.2; DB 7; Length 1257;
Best Local Similarity  64.1%; Pred. No. 0.24;
Matches  50; Conservative  0; Mismatches 28; Indels  0; Gaps  0;

Qy  671 CCGACAAGTCTCATTAACCGGTGCGGGTGGGTAAGCCTCGCATCAATGCTGACGGAGATATTTC 730
            ||||  | ||| ||  |  ||  | || || |||||  ||||| | || ||| |||| |||||
Db  253 CCGAGTCTGTCTCAAAAAATGGGTTTGGTGAGCCGTGCCATCAGCGCGCTGAAGGAGGCGTATT 194

Qy  731 TCCTGCAACAAGCACACAGG 748
            || | | || | ||  |||
Db  193 TCATCCAATCCGTAGAGG 176
```

```
RESULT 6
US-11-293-654-37
```

; Sequence 37, Application US/11293654
; Publication No. US20060122368A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; APPLICANT: Represented by the Secretary of the Dept. of Health and Human
; APPLICANT: Services
; APPLICANT: Patton, John T.
; APPLICANT: Taraporewala, Zenobia P.
; TITLE OF INVENTION: ANTIGEN DELIVERY PLATFORM
; FILE REFERENCE: 4239-69406-02
; CURRENT APPLICATION NUMBER: US/11/293,654
; CURRENT FILING DATE: 2005-12-02
; PRIOR APPLICATION NUMBER: US 60/633,036
; PRIOR FILING DATE: 2004-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin version 3.3
;
; SEQ ID NO 37
; LENGTH: 3431
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Vector
US-11-293-654-37

Query Match   66.7%;   Score 32.2;   DB 7;   Length 3431;
Best Local Similarity   66.7%;   Score 32.2;   DB 7;   Length 3431;
Matches   46;   Conservative   0;   Mismatches   23;   Indels   0;   Gaps   0;

Qy   1255 ACGCAAATCGTGAATGAAGCACGGCATACCGGGCCTGCGCAGTTTGAGATCTCATCACCATCAC 1314
          |||||||||||||||||||||||||||||||||||||||||||||
Db   85 ACAGAATTCATTAAAGAGAGAGAAAATTAACCATGGGAGGATCCAGATCTCATCACCATCAC 144

Qy   1315 CATCACTAA 1323
          ||||||||
Db   145 CATCACTAA 153

RESULT 7
US-10-449-902-8987
; Sequence 8987, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8987
; LENGTH: 1396
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK106335
; DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-8987

Query Match   58.3%;   Score 32;   DB 6;   Length 1396;
Best Local Similarity   58.3%;   Score 32;   DB 6;   Length 1396;
Matches   56;   Conservative   0;   Mismatches   40;   Indels   0;   Gaps   0;

Qy   679 GTCTCATTAACCGGTGCGGTAAGCCTGCCATCAAATGCTGACCGGAGAGATATTTCTCCTGCAA 738
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   456 GTCTCGTTCTCGGTCGCGGTCGCCCTTCTCGCCCGAAGCGGGAGGAGGAGCTCGTACTCCGAGAA 515

Qy   739 CAAGCACAGGGAAATGCCCGGAGCCGGGGGTGGGGGAAGG 774
          ||||||||||||||||||||||||||||||||||||||
Db   516 CCGGCATGGGGACAGGCGGGCGCGACGAATGGTGGGCGG 551

RESULT 8
US-10-449-902-6188/c
; Sequence 6188, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6188
; LENGTH: 1489
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK104957
; DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-6188

Query Match   51.9%;   Score 31.6;   DB 6;   Length 1489;
Best Local Similarity   51.9%;   Score 31.6;   DB 6;   Length 1489;
Matches   70;   Conservative   0;   Mismatches   65;   Indels   0;   Gaps   0;

Qy   213 GCGNGGTGGTGAGCTAAATGGCCTATCTCGGACATTACCAAGCGACCGTGTGGGTAGCCGA 272
          |||||||||||||||||||||||||||||||||||||||||||
Db   538 GCCGGGCAGTCGACGGCCACCCCCATGGGCGAAAGCTCACCCGCCGGCCGCCACGTCGACGC 479

Qy   273 CGGATTGCTCGGCGAAAAAGGGGCTGCCCCGCAGTCTCGGTCAGGTCGGGATTAATTGCTGCATGT 332
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   478 CGCCGAGCACCTTCTCCAGCTCGGCGGCGCTCACCACCATCAGGTGCGTCCTCCGGTGGTCGGTGT 419

Qy   333 CGACGAGCGCTACCCG 347
          |||||||||||||||
Db   418 CGTCCTCCTTCTCCG 404

RESULT 9
US-10-449-902-17204/c
; Sequence 17204, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 17204
; LENGTH: 1600
; TYPE: DNA
; ORGANISM: Oryza sativa

; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK067664
; DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-17204

Query Match   70; Conservative   0; Mismatches   65; Indels   0; Gaps   0;
Best Local Similarity   51.9%; Pred. No. 1;
Matches   70; Conservative   0; Mismatches   65; Indels   0; Gaps   0;

Qy   213 GCGNGGTGGTGAGCTAATCGGCCTATCTCGGACATTACCAAGCCACGTGCTCTGGTAGCCGA 272
Db   615 GCGGGCAGTCGACGCCACCCGCCCATGGCGAAGCTCACCCGCGCGGAGCCACGTCGACGC 556

Qy   273 CGGATTGCTGGCGAAAAAGGGCTTGCCCGCAGTTCTGGTCGGCGATTAATTGCTGATGT 332
Db   555 CGCCGAGCACACTTTCTCCAGCTCGGCGCTCACCACCATCAGGTCGTCGTCTCCGTGGTGGTGT 496

Qy   333 CGACGAGCGTACCCG 347
Db   495 CGTCGTCCTTCTCCG 481

RESULT 10
US-10-449-902-13863/c
; Sequence 13863, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 13863
; LENGTH: 3129
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK111211
; DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-13863

Query Match   47.7%; Pred. No. 1.8; DB 6; Length 3129;
Best Local Similarity   47.7%; Pred. No. 1.8; DB 6; Length 3129;
Matches   92; Conservative   0; Mismatches 101; Indels   0; Gaps   0;

Qy   999 CGGCGCACTGGAGCTCAACTGGACGCTTCCCGTCACCGGATAAACACGCCCCAGGTGG 1058
Db   2942 CGAAGTTCCGGAGCTCAGACTGCCGCTTCCATGGCTTCCGTTGCCCGTTGAGATGGGG 2883

Qy   1059 TGAACTGGTGTTTGAACGCTGGCGTCCGGCTAAGCGATTAAACAGCCAGTGGGATTCAGGTTTC 1118
Db   2882 TGCGTGGAAACGGCGGTTACCGAGTTGGGAAGCCAGAGGGTGAGTCGGGCTTCGTCGAAGCCATC 2823

Qy   1119 GCTGGTGTCTTCCAGACTTTACAGCAGATGCGGTTGATAAAACGCCCGCCTGTTCATTAAATACGCC 1178
Db   2822 GGTGAACCCGGAGAGTGTTGGGAGGCAATGGACGAAGGAACGGCGCGCGGGCCATTGCGGACGGGA 2763

Qy   1179 GCCCGAGGAGGTG 1191
Db   2762 ACCAGGAGTGCTG 2750

RESULT 11
US-10-601-319-1.rnpbn
; Sequence 12516, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 12516
; LENGTH: 2144
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK109864
; DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-12516

Query Match   48.6%; Pred. No. 1.7; DB 6; Length 2144;
Best Local Similarity   48.6%; Pred. No. 1.7; DB 6; Length 2144;
Matches   84; Conservative   0; Mismatches   89; Indels   0; Gaps   0;

Qy   211 CCGCGNGGTGGTGAGCTAATCGGCCTATCTCGGACATTACCAAGCCACGTGCTGGTAGCC 270
Db   1652 CGGCGGGGTGGGGCCACCAGCCTGGCTTCAGCCGGAGCGGGCGGGTACTCCATCCCTTCTTCTGTCCGCGG 1593

Qy   271 GACGGATTGCTGGCGAAAAAGGGCTTGCCCGCAGTTCTGGTCGGCGATTAATTGCTGAT 330
Db   1592 GTTGTAGTGGTTCGCCGCGGATCGCCACCTTCTTCCTTCAGCTTCGCGCCTTGCGGACGAA 1533

Qy   331 GTCGACGAGCGTACCCGTAAAACAGGCGGAAGCCTTCGCCGCCGCGGCTGGGCACC 383
Db   1532 GTCCACCCACGTCTCGTTGTTGCCGGACGCCGCGCGCGGCGGGCGTCCCAGC 1480

RESULT 12
US-10-449-902-10760/c
; Sequence 10760, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10760
; LENGTH: 1344
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK108108
; DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-10760

Query Match   53.8%; Pred. No. 1.5; DB 6; Length 1344;
Best Local Similarity   53.8%; Pred. No. 1.5;

Matches 64; Conservative   0; Mismatches 55; Indels   0; Gaps   0;

```
Qy   258 GCGTCTGGGTAGCCGACGGATTGCTGGGAAAAAGGCTGCCCGCAGTCTGGTCAGGTCGC 317
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   289 GCGTTCGGGTTAGTCGATNGCCCTACGGGCCGGTCCCGGTGGTTGGTTAGGGTAGT 230

Qy   318 GATTATTGTTGTGATGTCGACGGAGCGTTACCCGTTAAAACAGGCGGAACCTTCGGCGACCGGGGGC 376
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   229 TGTTGGTTGCCGTTGTCGACCGAGGGGAGTCGCCGAACACGAAGAAGGCCCGCTCCCGGGAC 171
```

RESULT 13
US-11-293-697-707/c
Sequence 707, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR APPLICATION FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 707
; LENGTH: 2791
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-707

Query Match   2.3%; Score 31; DB 7; Length 2791;
Best Local Similarity   55.0%;
Matches 61; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    59 CATTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAAGTGTGGTTGGTTGTCAGTGTCGTCATG 118
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1419 CCTTCGCAGTGAGTGTTACAGCTCTTTAAGGCGGGTGTGTCTGGAGTTGTTCATTCCTCTG 1360

Qy   119 GTGTGCGTGCTCCAACCAAGGCCACGCAACTGATNGCAGGAATGTCACCCCAG 169
             |||||||||||||||||||||||||||||||||||||||||||||
Db  1359 GTCGGGTTCGTGGTCTCGCTGGCCTCAGAAGTGAAGCTCGAGATCTTCCGGG 1309
```

RESULT 14
US-10-953-349-27515/c
Sequence 27515, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCONDED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
;
; SEQ ID NO 27515
; LENGTH: 1393
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-953-349-27515

Query Match   2.3%; Score 30.8; DB 6; Length 1393;
Best Local Similarity   47.8%; Pred. No. 1.8;
Matches 89; Conservative   0; Mismatches 97; Indels   0; Gaps   0;

```
Qy   603 TAAACGTGAGAAAACAGGACGAAAGCTGTTCATTACGGAGCGATTACCATCGGAAACTCAA 662
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1167 TAACCTTTAGGTAGTACATCTTGCAGGTTCTCCAACAGTTAAGCTTAACAGCTGACAGTCA 1108

Qy   663 GGTGAGCGCCGACAAATGTCTCATTAACCGGTCGGGTAAGCCTCGCATCAAATGCTGACCGA 722
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1167                                                           663
```

RESULT 15
US-10-449-902-11476/c
Sequence 11476, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 11476
; LENGTH: 1461
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK108824
; DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-11476

Query Match   2.3%; Score 30.8; DB 6; Length 1461;
Best Local Similarity   53.3%; Pred. No. 1.9;
Matches 65; Conservative   0; Mismatches 57; Indels   0; Gaps   0;

```
Qy   253 CGCCAGCGTCTGGTTAGCCGACGGATTGCTGTGGGAAAAAGGGGCTTCCCCGCCAGTCTCGGTCAG 312
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1072 CGGGCGACGGCGCGGGTTGACAGCGAGAGCACTTGTGCCTGATCTCGCCGTGGGAGTGGGGTGAG 1013

Qy   313 GTCGCGGATTATTTGCTGAATGTCGACGAGCGTTACCCGTAAAACAGGCGGAAGGCCTTCGGCCGCC 372
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1012 CACGCCGACTTTGCTCATGTTGACGAAGGAGGACCGATGAAGTCGCGAAGAACTCGTCGTC 953

Qy   373 GG 374
             ||
Db   952 GG 951
```

Search completed: June 14, 2006, 14:23:15
Job time : 41.2281 secs

THIS PAGE LEFT BLANK

OM nucleic - nucleic search, using sw model

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

Run on: June 14, 2006, 10:29:05 ; Search time 1615.16 Seconds
(without alignments)
10064.996 Million cell updates/sec

Title: US-10-601-319-1
Perfect score: 1322
Sequence: 1 atgaaagcgatcttaatccc..........atcaccaccaccactaa 1323

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 18892170 seqs, 614307638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main:*
1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09_PUBCOMB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US0A_PUBCOMB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US0B_PUBCOMB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US0D_PUBCOMB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US0E_PUBCOMB.seq:*
11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US0F_PUBCOMB.seq:*
12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US0G_PUBCOMB.seq:*
13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US1A_PUBCOMB.seq:*
14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US1B_PUBCOMB.seq:*
15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US1C_PUBCOMB.seq:*
16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US1D_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	1322	100.0	1323	3	US-09-777-566A-1	Sequence 1, Appli
2	1322	100.0	1323	3	US-09-866-379-1	Sequence 1, Appli
3	1322	100.0	1323	6	US-10-034-985-1	Sequence 1, Appli
4	1322	100.0	1323	7	US-10-430-356-1	Sequence 1, Appli
5	1322	100.0	1323	8	US-10-601-319-1	Sequence 1, Appli
6	1322	100.0	1323	10	US-10-993-115-1	Sequence 1, Appli
7	1287.8	97.4	1901	6	US-09-866-379-7	Sequence 7, Appli
8	1287.8	97.4	1901	6	US-10-155-660-3	Sequence 3, Appli
9	1287.8	97.4	1901	6	US-10-601-319-7	Sequence 7, Appli
10	1287.8	97.4	1901	10	US-10-933-115-7	Sequence 7, Appli
11	1287.8	97.4	1901	14	US-11-056-354-3	Sequence 3, Appli
12	1287	97.4	1299	3	US-10-282-122A-7167	Sequence 7167, Ap
13	1286.2	97.3	1901	3	US-09-866-379-5	Sequence 5, Appli
14	1286.2	97.3	1901	9	US-09-866-379-9	Sequence 9, Appli
15	1284.6	97.2	1901	8	US-10-601-319-5	Sequence 5, Appli
16	1284.6	97.2	1901	10	US-10-933-115-5	Sequence 5, Appli
17	1283	97.0	1901	3	US-09-866-379-6	Sequence 6, Appli
18	1283	97.0	1901	8	US-10-601-319-6	Sequence 6, Appli
19	1283	97.0	1901	10	US-10-933-115-6	Sequence 6, Appli
20	1308	95.7	1265	6	US-10-156-660-1	Sequence 1, Appli
21	1308	95.7	1265	8	US-10-601-319-9	Sequence 9, Appli
22	1308	95.6	1265	10	US-10-933-115-9	Sequence 9, Appli
23	1489	95.6	1263.8	10	US-10-266-041-9	Sequence 9, Appli
24	1489	95.6	1263.8	7	US-10-284-962-1	Sequence 1, Appli
25	1486	95.6	1489	7	US-11-018-709-2	Sequence 2, Appli
26	1296	94.9	1255	14	US-11-056-354-1	Sequence 1, Appli
27	1486	94.7	1251.8	7	US-11-074-522-14	Sequence 14, Appl
28	728.4	94.7	1251.8	13	US-11-334-672-4	Sequence 4, Appli
29	689.4	52.1	1737	10	US-11-450-763-25689	Sequence 25689, A
30	666.4	50.4	1281	7	US-11-450-763-24247	Sequence 24247, A
31	666.4	50.4	1281	7	US-10-450-763-30120	Sequence 30120, A
32	399.4	30.2	4050	10	US-10-282-122A-31608	Sequence 31608, A
33	399.4	30.2	466	8	US-10-282-122A-41608	Sequence 41608, A
34	399	29.5	466	8	US-10-021-723A-3	Sequence 3, Appli
35	390	29.5	1264	10	US-10-450-763-24247	Sequence 24247, A
36	390	29.5	1264	10	US-10-450-763-30120	Sequence 30120, A
37	222.8	16.9	1326	8	US-10-282-122A-3181	Sequence 3181, Ap
38	218.8	16.3	1325	6	US-10-021-723A-3	Sequence 3, Appli
39	215.6	16.3	1325	6	US-10-021-723A-1	Sequence 1, Appli
40	206.6	15.6	1326	10	US-10-450-763-25687	Sequence 25687, A
41	199	15.1	356	10	US-10-021-723A-11	Sequence 11, Appl
42	164.2	12.4	254	7	US-10-317-444-301	Sequence 301, App
43	164.2	12.4	254	7	US-10-317-444-302	Sequence 302, App
44	164.2	12.4	254	7	US-10-317-444-303	Sequence 303, App
45	164.2	12.4	254	7	US-10-317-444-304	Sequence 304, App

ALIGNMENTS

RESULT 1
US-09-777-566A-1
Sequence 1, Application US/09777566A
Patent No. US20010055788A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: SHORT, Jay
APPLICANT: KRETZ, Keith
TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
FILE REFERENCE: DIVER1370-6
CURRENT APPLICATION NUMBER: US/09/777,566A
CURRENT FILING DATE:
PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn version 3.0

SEQ ID NO 1
LENGTH: 1323
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1320)
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(1323)
OTHER INFORMATION: n is any nucleotide
US-09-777-566A-1

Query Match 100.0%; Score 1322; DB 3; Length 1323;
Best Local Similarity 100.0%;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy  1 ATGAAAGCGATCTTAATCCCATTTTAATCTTTCTGATTCCGTTAACCCCGCAAATCTGCA 60
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy      1 ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTCGATTCCGTTAACCCCGCAAATCTGCA 60

Qy     61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGGTCATTGTCAGTCGTCATGGT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGGTCATTGTCAGTCGTCATGGT 120

Qy    121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGAGCACAGATGCCATGGCCA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGAGCACAGATGCCATGGCCA 180

Qy    181 ACCTGGCCGGGTAAAAACTGGGTTGGCTGACACCCGAGGCTAATCGCCTATCTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACCTGGCCGGGTAAAAACTGGGTTGGCTGACACCCGAGGCTAATCGCCTATCTC 240

Qy    241 GGACATTACCAAACGCCAGCGTCTGGTAGCCGACGGAGCTGCCCG 300
          ||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGACATTACCAAACGCCAGCGTCTGGTAGCCGACGGAGCTGCCCG 300

Qy    301 CAGTCTCGGTCAGGTCGGCCGGATTATTGCTGATGTGCGAAAAAGGGCTGCCG 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGTCTCGGTCAGGTCGGCCGGATTATTGCTGATGTGCGAAAAAGGGCTGCCG 360

Qy    361 GCCTTCGCCGCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCCAGACG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCTTCGCCGCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCCAGACG 420

Qy    421 TCCAGTCCCGATCCGTTATTTAATCCCTCTAAAAAACTGGCGTTTGCCAACTGGATAACGG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCCAGTCCCGATCCGTTATTTAATCCCTCTAAAAAACTGGCGTTTGCCAACTGGATAACGG 480

Qy    481 AACGTGACTGAGCGCGATCCTCAGCAGGGCGCAGAGGGTCAATTGCTGACTTTACCGGGCAT 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AACGTGACTGAGCGCGATCCTCAGCAGGGCGCAGAGGGTCAATTGCTGACTTTACCGGGCAT 540

Qy    541 CGGCAAACGGCGTTTCGGGAACTGGAACGGGTTGCTTAATTTTCCGCAATCAAACTTGTGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CGGCAAACGGCGTTTCGGGAACTGGAACGGGTTGCTTAATTTTCCGCAATCAAACTTGTGC 600

Qy    601 CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCATTACGCAGGCCATTACCATCGAATAACGCG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCATTACGCAGGCCATTACCATCGGAACTC 660

Qy    661 AAGGTGAGCGCCGACAAATGTCTCATTAACCGGTGCGGTTAAGCCTCGCATCAAATGCTGACG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AAGGTGAGCGCCGACAAATGTCTCATTAACCGGTGCGGTTAAGCCTCGCATCAAATGCTGACG 720

Qy    721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCTCGGAAGGATCACC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCTCGGAAGGATCACC 780

Qy    781 GATTCACACCAGTGGGACACACCTTGCTAAGTTGCATAACGCGCAATTTTATTTTGCTACAA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GATTCACACCAGTGGGACACACCTTGCTAAGTTGCATAACGCGCAATTTTATTTTGCTACAA 840

Qy    841 CGCACGCGCAGAGGTTGCCGCGCCGCCACCCGGTTATTGGATTTGGATCATTGGGCAGCG 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CGCACGCGCAGAGGTTGCCGCGCCGCCACCCGGTTATTGGATTTGGATCATTGGGCAGCG 900

Qy    901 TTGACGCGCCCATCCACCGCCAAAAACAGGCGTATGGGTGTGACATTACCCAGTACTG 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTGACGCGCCCATCCACCGCCAAAAACAGGCGTATGGGTGTGACATTACCCAGTACTG 960

Qy    961 TTTATTTGCCGGACACGATACTAATCTGCAAAATCTCGGCGGCGCACTGGAGCTCAACTGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTTATTTGCCGGACACGATACTAATCTGCAAAATCTCGGCGGCGCACTGGAGCTCAACTGG 1020

Qy   1021 ACGCTTCCCGGTCAGCCGGATAAACACGCCGCGGTCAGGTGGTGAACTGGGTGTTTGAAACGGCTGG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGCTTCCCGGTCAGCCGGATAAACACGCCGCGGTCAGGTGGTGAACTGGGTGTTTGAAACGGCTGG 1080

Qy   1081 CGTCGGCTAAGCGGATAACAGCCAGTGGGATTCAGGGTTTGGCTGGTGCTTCTTCCAGAGACTTTACAG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CGTCGGCTAAGCGGATAACAGCCAGTGGGATTCAGGGTTTGGCTGGTGCTTCTTCCAGAGACTTTACAG 1140
```

Query Match   100.0%; Score 1322; DB 3; Length 1323;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTCGATTCCGTTAACCCCGCAAATCTGCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTCGATTCCGTTAACCCCGCAAATCTGCA 60

Qy     61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGGTCATTGTCAGTCGTCATGGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGGTCATTGTCAGTCGTCATGGT 120

Qy    121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGAGCACAGATGCGCATGGCCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGAGCACAGATGCGCATGGCCA 180
```

```
Qy   1141 CAGATGCGTGGATGAATAAAAAACGCACGCCTGCCGTGTCATTAAATATACGGCACCCGCCCGGGAGAGGTGAAAAACTGACC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAGATGCGTGGATGAATAAAAAACGCACGCCTGCCGTGTCATTAAATATACGGCACCCGCCCGGGAGAGGTGAAAAACTGACC 1200

Qy   1201 CTGGCAGGATGTGAAGAGGCGAAATGCGGCAGGGCCATGTGNGTTCGTTGGCAGGTTTTACCGCAA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGCAGGATGTGAAGAGGCGAAATGCGGCAGGGCCATGTGNGTTCGTTGGCAGGTTTTACCGCAA 1260

Qy   1261 ATCGTGAATGAAGGCACGCCGATACCCGGGCGGTGCCAGTTTGAGATCTCATCACCATCAC 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAGGCACGCCGATACCCGGGCGGTGCCAGTTTGAGATCTCATCACCATCAC 1320

Qy   1321 TAA 1323
          |||
Db   1321 TAA 1323
```

```
RESULT 3
US-10-034-985-1
; Sequence 1, Application US/10034985
; Publication No. US20030049815A1
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/034,985
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1323
;   TYPE: DNA
;   ORGANISM: Escherichia coli
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(1320)
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1323)
;   OTHER INFORMATION: n = A,T,C or G
US-10-034-985-1

Query Match            100.0%;   Score 1322;   DB 6;   Length 1323;
Best Local Similarity  100.0%;
Matches 1323;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```
Qy    1 ATGAAAGCGGATCTTAATCCCATTTTTACTCTTCTGATTCCGTTAACCCTGCCAATCTGCA   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAGCGGATCTTAATCCCATTTTTACTCTTCTGATTCCGTTAACCCTGCCAATCTGCA   60

Qy   61 TTCGCTCAGAGTGAGCCCGGAGCTGAAGCTGGAAAAGTGTGGGTGATTGTTCGTCGTCATGGT  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCCGGAGCTGAAGCTGGAAAAGTGTGGGTGATTGTTCGTCGTCATGGT  120

Qy  121 GTGGGTGCTTCCAACCAAGGCCAGCCAACTGATGCGGAGGATGCA  180
        |||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTGGGTGCTTCCAACCAAGGCCAGCCAACTGATGCGGAGGATGCA  180

Qy  181 ACCTGGCCCGGTAAAACTGGGTTGGCTGACACCCGCGNNGGCTATACTCTC  240
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ACCTGGCCCGGTAAAACTGGGTTGGCTGACACCCGCGNNGGCTATACTCTC  240

Qy  241 GGACATTACCAACGCCACGCGTCTGGGTAGCCGACGAGGATTGCTCGGGCTGCCCG  300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GGACATTACCAACGCCACGCGTCTGGGTAGCCGACGAGGATTGCTCGGGCTGCCCG  300

Qy  301 CAGTCTGGTTCAGGTCGCGGATTATTGCTGAGTGCGACGAGCGTAAAACAGGCGGAA  360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGTCTGGTTCAGGTCGCGGATTATTGCTGAGTGCGACGAGCGTAAAACAGGCGGAA  360

Qy  361 GCCTTCGCGCGCCCGGGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG  420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GCCTTCGCGCGCCCGGGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG  420
```

```
Qy 1261 ATCGTGAATGAAGCACGCATACCGGCGTGCGTGCAGTTTGAGATCTCATCACCATCAC  1320
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 ATCGTGAATGAAGCACGCATACCGGCGTGCGTGCAGTTTGAGATCTCATCACCATCAC  1320

Qy 1321 TAA  1323
        |||
Db 1321 TAA  1323
```

```
;   FILE REFERENCE: 09010/029003
;   CURRENT APPLICATION NUMBER: US/10/430,356
;   CURRENT FILING DATE: 2003-05-05
;   PRIOR APPLICATION NUMBER: US/09/580,515
;   PRIOR FILING DATE: 1999-05-25
;   PRIOR APPLICATION NUMBER: 03/318,528
;   PRIOR FILING DATE: 1999-05-25
;   PRIOR APPLICATION NUMBER: 09/291,931
;   PRIOR FILING DATE: 1999-04-13
;   PRIOR APPLICATION NUMBER: 08/910,798
;   PRIOR FILING DATE: 1997-08-13
;   PRIOR APPLICATION NUMBER: 09/259,214
;   PRIOR FILING DATE: 1999-03-01
;   NUMBER OF SEQ ID NOS: 4
;   SOFTWARE: FastSEQ for Windows Version 4.0
;
;   SEQ ID NO 1
;   LENGTH: 1323
;   TYPE: DNA
;   ORGANISM: Escherichia coli
;   FEATURE:
;   NAME/KEY:
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(1320)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1323)
;   OTHER INFORMATION: n = A,T,C or G
US-10-430-356-1

Query Match         100.0%;   Score 1322;   DB 7;   Length 1323;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1323;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;
```

RESULT 4
US-10-430-356-1
; Sequence 1, Application US/10430356
; Publication No. US20030232041A1
; Publication
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
;
·

```
RESULT 5
US-10-601-319-1
  Sequence 1, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING THEM
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; CURRENT FILING DATE:   2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
```

```
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
;
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 216
; OTHER INFORMATION: n = A,T,C or G
US-10-601-319-1

Query Match             100.0%;  Score 1322;  DB 8;  Length 1323;
Best Local Similarity   100.0%;
Matches 1323;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCGCCAATCTGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCGCCAATCTGCA 60

Qy         61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGCTGGAAAAGTGTGGTGATTGTCAGTCGTCATGGT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGCTGGAAAAGTGTGGTGATTGTCAGTCGTCATGGT 120

Qy        121 GTGGCTGCTCCAACCAAGGCCACCAGCCAACTGATGATGCAGGATGTCACCCCAGACGCA 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTGGCTGCTCCAACCAAGGCCACCAGCCAACTGATGATGCAGGATGTCACCCCAGACGCA 180

Qy        181 ACCTGGCCGGTAAAAACTGGGTTGGCCTAGCACCGCGGGTGGTGGAGCTAATGCCTACTC 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCTGGCCGGTAAAAACTGGGTTGGCCTAGCACCGCGGGTGGTGGAGCTAATGCCTACTC 240

Qy        241 GGACATTACCAAGCCCAGGGTCTCGGTAGCCGACGGCGGATTGCTGGCCGGAAAAAAAGGGCTGCCCG 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGACATTACCAAGCCCAGGGTCTCGGTAGCCGACGGCGGATTGCTGGCCGGAAAAAAAGGGCTGCCCG 300

Qy        301 CAGTCTGGTCGCGATTAATTGCTGCGACGACGTACCCGGTAAAAACAGGGCGAA 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGTCTGGTCGCGATTAATTGCTGCGACGACGTACCCGGTAAAAACAGGGCGAA 360

Qy        361 GCCTTCGCCGCGCCGGGCTGGCCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTTCGCCGCGCCGGGCTGGCCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420

Qy        421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAAACTGGCGTTTGCCAACTGGATAACGCG 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAAACTGGCGTTTGCCAACTGGATAACGCG 480

Qy        481 AACGTGACTGACCCGGATCCTCAGCACGGGCAGGAGGGGTCAATTGCTGACTTTACGGGCAT 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACGTGACTGACCCGGATCCTCAGCACGGGCAGGAGGGGTCAATTGCTGACTTTACGGGCAT 540

Qy        541 CGGCAAACGGCGTTTCGGGAACTGGACGGGTGCTTAATTTCCGGCAATCAAAACTTGTGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGGCAAACGGCGTTTCGGGAACTGGACGGGTGCTTAATTTCCGGCAATCAAAACTTGTGC 600

Qy        601 CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCATTAACGGCAGGCCATTACCAATCGGGAACTC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCATTAACGGCAGGCCATTACCAATCGGGAACTC 660

Qy        661 AAGGTGACGCGCCACAACAAGTCTCATTAACCGGTGCGCATCAATGCTGACG 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy        601 CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCATTAACGGCAGGCCATTACCAATCGGGAACTC 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGGTGAGCGGCCGGACAATGTCTCATTATTACCGGTGCGGTAAGCCTTCGGCATCAATGCTGACG 660

Qy        661 AAGGTGACGCGCCGACAATGTCTCATTATTACCGGTGCGGTAAGCCTTCGGCATCAATGCTGACG 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGGTGAGCGGCCGGACAATGTCTCATTAACCGGTGCGGGTAAGCCTTCGGCATCAATGCTGACG 720

Qy        721 GAGATATTTCTCTGCGACAACAAGCACACAGGGAATGCCGGAGCCGGGGGGAAGGATCACC 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGATATTTCTCTGCGACAACAAGCACACAGGGAATGCCGGAGCCGGGGGGAAGGATCACC 780

Qy        781 GATTCACCAGTGGAACACCTTGCTAACGGCGCAAATTTATTTATTGCTACAA 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GATTCACCAGTGGAACACCTTGCTAACGGCGCAAATTTATTTATTGCTACAA 840

Qy        841 CGCACGCCAGAGGTTGCCCGCGCGCCACCCCGTTATTGGATTTGATTCATCAGCACGCG 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CGCACGCCAGAGGTTGCCCGCGCGCCACCCCGTTATTGGATTTGATTCATCAGCACGCG 900

Qy        901 TTGACGCGCCCATCCACCGCAAAAAACAGGGGTATGGTGTGACATTACCCACTTCAGTACTG 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTGACGCGCCCATCCACCGCAAAAAACAGGGGTATGGTGTGACATTACCCACTTCAGTACTG 960

Qy        961 TTTATTGCCGGACACGATACTAATCTGGCAAAATCTCGCGGCGACTGCTCAACTGG 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTTATTGCCGGACACGATACTAATCTGGCAAAATCTCGCGGCGACTGCTCAACTGG 1020

Qy       1021 ACGCTTCCCGGTCAGCCGGATAACACGCGCCGCCCAGGTGGTGAACTGGGTGTTGTTGAACGGCTGG 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACGCTTCCCGGTCAGCCGGATAACACGCGCCGCCCAGGTGGTGAACTGGGTGTTGTTGAACGGCTGG 1080

Qy       1081 CGTCGGCTAAGCGGATAACAGCCAGTGGATTCAGGTTCGCTGGTTCTTCCAGACTTTACAG 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CGTCGGCTAAGCGGATAACAGCCAGTGGATTCAGGTTCGCTGGTTCTTCCAGACTTTACAG 1140

Qy       1141 CAGATGCCTGATAAAAACCGCGCTGTCATTAAATACGGCCCCGAGAGGTGAAAACTGACC 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CAGATGCCTGATAAAAACCGCGCTGTCATTAAATACGGCCCCGAGAGGTGAAAACTGACC 1200

Qy       1201 CTGGCACGGATTGTGAAAGAGCGAAATGCCGAGGGCCATGTGTTCGTTGGCAGGTTTACGGCAA 1260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CTGGCACGGATTGTGAAAGAGCGAAATGCCGAGGGCCATGTGTTCGTTGGCAGGTTTACGGCAA 1260

Qy       1261 ATCGTGAATGAAGCACGGCATACCGGGCGGTGCAGTTTGAGAGATCTCATCACCATCACCATCAC 1320
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ATCGTGAATGAAGCACGGCATACCGGGCGGTGCAGTTTGAGAGATCTCATCACCATCACCATCAC 1320

Qy       1321 TAA 1323
              |||
Db       1321 TAA 1323
```

RESULT 6
US-10-933-115-1
; Sequence 1, Application US/10933115
; Publication No. US20050281792A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O'Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 5644620001822
; CURRENT APPLICATION NUMBER: US/10/933,115
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214

; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 216
; OTHER INFORMATION: n = A,T,C or G
US-10-933-115-1

Query Match   100.0%;   Score 1322;   DB 10;   Length 1323;
Best Local Similarity 100.0%;   Pred. No. 0;
Matches 1323;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 ATGAAAGCGATCTTAARTCCA... 60
Db   1 ATGAAAGCGATCTTAARTCCA... 60

Qy   61 TTCGCTCAGAGTGAGCCGGA... 120
Db   61 TTCGCTCAGAGTGAGCCGGA... 120

Qy 121 GTGCGTGCTTCCAACCAAGG... 180
Db 121 GTGCGTGCTTCCAACCAAGG... 180

Qy 181 ACCTGGCGCCGGTAAAACTG... 240
Db 181 ACCTGGCGCCGGTAAAACTG... 240

Qy 241 GGACATTACCAACGCCAACG... 300
Db 241 GGACATTACCAACGCCAACG... 300

Qy 301 CAGTCTGGTCAGGTCCGGAT... 360
Db 301 CAGTCTGGTCAGGTCCGGAT... 360

Qy 361 GCCTTCGGCCGCGCCGGGTG... 420
Db 361 GCCTTCGGCCGCGCCGGGTG... 420

Qy 421 TCCAGTCCCGATCCGTTGAT... 480
Db 421 TCCAGTCCCGATCCGTTGAT... 480

Qy 481 AACGTGACTGACGGCGATCC... 540
Db 481 AACGTGACTGACGGCGATCC... 540

Qy 541 CGGCAAACGGGCGTTTCGCG... 600
Db 541 CGGCAAACGGGCGTTTCGCG... 600

Qy 601 CTTAAACGTGAGAAAACAGG... 660
Db 601 CTTAAACGTGAGAAAACAGG... 660

Qy 661 AAGGTGAGCGCGCGACAATG... 720
Db 661 AAGGTGAGCGCGCGACAATG... 720

Qy 721 GAGATATTTCTCCTGCAACA... 780
Db 721 GAGATATTTCTCCTGCAACA... 780

Db 661 AAGGTGAGCGCCCGACAATGTCTCATTAACCGGTGCGTGCGTAAGCCTCGCATCAAATGCTACG 720
Qy 721 GAGATATTTCTCCTGCAACAAGCACACAGGGAATGCCCGGAGCCGGGGTGGGGAGGGATCACC 780
Db 721 GAGATATTTCTCCTGCAACAAGCACACAGGGAATGCCCGGAGCCGGGGTGGGGAGGGATCACC 780

Qy 781 GATTCACACCAGTGGAACACCCTTGCTTAAGTTTTGCATAAACGGGCAATTTTATTTGCTACAA 840
Db 781 GATTCACACCAGTGGAACACCCTTGCTTAAGTTTTGCATAAACGGGCAATTTTATTTGCTACAA 840

Qy 841 CGCACGCCAGAGGTTGCCCGCGGCCGCCACCCGGTTATTAACGGCCAATTTATTTGCTACGC 900
Db 841 CGCACGCCAGAGGTTGCCCGCGGCCGCCACCCGGTTATTAACGGCCAATTTATTTGCTACGC 900

Qy 901 TTGACGCCCCCATCCACCGCAAAAAACAGGCGTATGGGTGTGACATTACCCACTTCAGTACTG 960
Db 901 TTGACGCCCCCATCCACCGCAAAAAACAGGCGTATGGGTGTGACATTACCCACTTCAGTACTG 960

Qy 961 TTTATTGCCGGACACGATACTAATCTCGGCAAATCTCGGGCGCACTGGGAGCTCACTGG 1020
Db 961 TTTATTGCCGGACACGATACTAATCTCGGCAAATCTCGGGCGCACTGGGAGCTCACTGG 1020

Qy 1021 ACGCTTCCCGGTCAGCCGGATAACACGCCCGGATAACACGCCCGGATGGTGGTGAACGGCTGG 1080
Db 1021 ACGCTTCCCGGTCAGCCGGATAACACGCCCGGATAACACGCCCGGATGGTGGTGAACGGCTGG 1080

Qy 1081 CGTCGGCTAAGGGATAACAGCGCCAGTGGATTCAGGTTTCGCTGGTCTCTTCCAGACTTTACAG 1140
Db 1081 CGTCGGCTAAGGGATAACAGCGCCAGTGGATTCAGGTTTCGCTGGTCTCTTCCAGACTTTACAG 1140

Qy 1141 CAGAGCGTGATAAAACGCCCCTGTCATTAAATACCGGCGAGAGGTGAAAACTGACC 1200
Db 1141 CAGAGCGTGATAAAACGCCCCTGTCATTAAATACCGGCGAGAGGTGAAAACTGACC 1200

Qy 1201 CTGGCAGGATGTGAAGAGCGAAAATGCCCAGGGGCATGTGTTCGTTGGCAGGTTTAGGCAA 1260
Db 1201 CTGGCAGGATGTGAAGAGCGAAAATGCCCAGGGGCATGTGTTCGTTGGCAGGTTTAGGCAA 1260

Qy 1261 ATCGTGAAATGAAGCACGCCATACCGGCGTGCAGTTTGAGATCTTCATCACCATCACCATCAC 1320
Db 1261 ATCGTGAAATGAAGCACGCCATACCGGCGTGCAGTTTGAGATCTTCATCACCATCACCATCAC 1320

Qy 1321 TAA 1323
Db 1321 TAA 1323

```
RESULT 7
US-09-866-379-7
Sequence 7, Application US/09866379
Patent No. US20020136754A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: SHORT, Jay
APPLICANT: KRETZ, Keith
APPLICANT: GRAY, Kevin
APPLICANT: BARTON, Nelson
APPLICANT: GARRETT, James
APPLICANT: O'DONOGHUE, Eileen
TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
FILE REFERENCE: DIVER1370-7
CURRENT APPLICATION NUMBER: US/09/866,379
CURRENT FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 09/580,515
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentin version 3.1
SEQ ID NO 7
LENGTH: 1901
TYPE: DNA
```

```
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(1901)
OTHER INFORMATION: n is any nucleotide
US-09-866-379-7

Query Match   97.4%;   Score 1287.8;   DB 3;   Length 1901;
Best Local Similarity   99.5%;   Pred. No. 0;
Matches 1293; Conservative   0;   Mismatches   7;   Indels   0;   Gaps   0;
```

```
Qy    781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCGCAAATTTTATTTGCTACAA   840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCGTAA           60

Qy    841 CGCACGCCCAGAGAGGTTGCCCGCACGCGCGGCCACCCGGTTATTGGATTTGATCATGGCACG   900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCGCCTCAGAGTGAAGCCCGAGAGCTGAAGCTGGAAAAGTGTGGTGATTGTCAGTCGTCATGGT   120

Qy    901 TTGACGCCCCATCCACCGCAAAAAACAGGGGTATGGGTGTGACATTACCCACTTCAGTACTG   960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGCGTGCTTCCAACCAAGGCCAGGCCAACTGAGTCAGATTGTCACCCGAGGCCA          180

Qy    961 TTTATTGCCGGACACGGATACTAATTCTGGCAAAATCTCGGCCGGACCTCAACTGG   1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188 TTCCGCTCAGAGTGAGCCCGAGCTGAAGCTGGAAAAGTGTGGTGATTGTCAGT          247

Qy   1021 ACGCTTCCCGGTCAGCCCGGATAACACGCGCCCAGGTGGTGAACTGGTGTTTGAACGCTGG   1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248 TTCCGCTCAGATGAGCCCGAGCTGAAGCTGGAAAAGTGTGGTGATTGTCAGTCGTCATGGT          307

Qy   1081 CGTTCGGGCTAAGCCGATAACAGCCAGTTGGAGTTCGGCTGTTCGCTGGCTTTACAG   1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308 GTGCGTGCTTCCAACCAAGGCCAGGCCAACTGAGTCAGATTGTCACCCGAGGCCA          367

Qy   1141 CAGATGCCTGATAAAAACGGCGCTGTCATTAAATACGGCGCCGCCGGAGGAGGTGAAAACTGACC   1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    368 ACCTGGCCGGTAAAAACTGGTTGGTTGGCTGACAGCCGGGTTGTGGCTAATCGCGCCTATCTC          427

Qy   1201 CTGGCAGGATGTGAAGGACGGAAAATGCGCAGGGGCAGTGTTGTTCGTGTTGGGCAGGTTTTTACGCAA   1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428 GGACATTACCAACGGCCAGGTCTCGGTAGCCGCACCGGATGCCGCAAAAAGGGGCTGCCCG          487

Qy   1261 ATCGTGAATGAAGGATGCCGAATACCGGCGGTTGCCAGTTTGAGATCTCATCACCACCATCAC   1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488 CAGTCTGGTCAGGTCGGCGATTATTGCTGATGTCGACGACGGGTACCCGGTAAAACAGGCGAA          547

Qy   1321 TAA   1323
          |||
Db    548 GCCTTCGCGCGCCGCGGCTGGCACCTGACCTGTGCAATAACCGTACATACCCAGGGACGATACG          607
```

```
Db   1088 TTGACCCCCATCCACGCGATAACGGCGTATGGTGGTGACATTACCCACTTCAGTGCTG 1147

Qy    961 TTTATTGCCGGACACGATACTAATCTTGGCAAATCTCGGCGGCCACTGGAGCTCAACTGG 1020
Db   1148 TTTATCGCCGGACACGATACTAATCTCGGCAAATCTCGGCGGCCACTGGAGCTCAACTGG 1207

Qy   1021 ACGCTTCCCGGTCAGCCGGATAACACGCCCGCCAGGTGGGTGAACTCGGTGTTTCAGTTGG 1080
Db   1208 ACGCTTCCCGGTCAGCCGGATAACACCCCCGCCAGGTGGGTGAACTCGGTGTTTCAGTGGT 1267

Qy   1081 CGTCGGCTTAAGCGATAACAGCCACGTGGATTCAGGTTTCGCTGGTCTTCCGATTTACAG 1140
Db   1268 CGTCGGCTTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCGAGACTTTACAG 1327

Qy   1141 CAGATGGCTGATAAAAAAAACGGCCGGCGTCATTAAAATAACGGCCGCCCGGGAGAGGTGAAAACTGACC 1200
Db   1328 CAGATGGGTGATAAAAAAAACGGCCCGTGTCATTAAATAACGGCCGCCCGGGAGAGGTGAAAACTGACC 1387

Qy   1201 CTGGCAGGGATGTGAAAGAGCGAAAATGCCCATGTTCTTCCTTGGCGCAGGTTTTTACGGCAA 1260
Db   1388 CTGGCAGGGATGTGAAAGAGCGAAAATGCCCAGGGCTCGGGTTAGGCAGGTTTTTACGGCAA 1447

Qy   1261 ATCGTGAAATGAAGCACGCCATACCGGCGGCGTGCAGTTTGTGAGAT 1300
Db   1448 ATCGTGAAATGAAGCACGCCATACCGGGCGGCGTGCAGTTTGTAAT 1487
```

```
                    Matches 1293;  Conservative  0;  Mismatches  7;  Indels  0;  Gaps  0;

Qy    361 GCCTTCGCGCGCCGGGGCGTTGGCACCTGACTGTGCAATAACGTACATACCAGGCAGATACG 420
Db    548 GCCTTCGCGCGCCGGGGCGTTGGCACCTGACTGTGCAATAACGTACATACCAGGCAGATACG 607

Qy    421 TCCAGTCCGCGATCCGGTTGATTTATTAATCCTTCTAAAAAACTGGGATTTGCCAAACTGAATAACGACG 480
Db    608 TCCAGTCCGCGATCCGGTTATTTATTTAATCCTTCTAAAAAACTGGGGATTTGCCAAACTGAATAACGACG 667

Qy    481 AACGTGACTGAGGCGGATCCTCAGCGACGGCAGGGGTCAATTGCTGACTTTACCGGCAT 540
Db    668 AACGTGACTGAGGCGGATCCTCAGCGACGGCAGGGGTCAATTGCTGACTTTACCGGCAT 727

Qy    541 CGGCAAACGGCGTTCGGGAACGGCAGGGTGCTTAATTTCCGCAATCAAAACTTGTGC 600
Db    728 CGGCAAACGGCGTTCGGGAACTGGGAACGGGCGTTAATTTCCGCAATCAAAACTTGTGC 787

Qy    601 CTTAAACGTGACAAAACAGGACGAAAAGCTGTTCATTAACGCAGGCCATTACCATCGGGAACTC 660
Db    788 CTTAAACGTGAGAAACAGGACGAAAAGCTGTTCATTAACGGCAGGCATTACCATCGGGAACTC 847

Qy    661 AAGGTGAGCGCGCCGACAATGTCTCATTAACCGGTGCCGTAAGCCTCGCCATCAATGCTCACG 720
Db    848 AAGGTGAGCGCGCCGACAATGTCTCATTAACCGGTGCCGTAAGCCTCGCCATCAATGCTCACG 907

Qy    721 GAGATATTTCTCCTGCAACAACAGGAAATGCCCGGACGGGGGGAGGGGATCACC 780
Db    908 GAGATATTTCTCCTGCAACAACAGGAAATGCCCGGACGGGGGTGGGGAATCACC 967

Qy    781 GATTCACCACCAGTGGGAACACCTTGCTTAAGTTTGCATAAACGGGCAATTTTATTTGTCTACA 840
Db    968 GATTCACCACCAGTGGGAACACCTTGCTTAAGTTTGCATAAACGGGCCAATTTTATTTGTCTACA 1027

Qy    841 CGCACGCCAGAGGTTGCCCGCGCCGCCACCCCGTTATTGGATTTGATCAAGACACAGCG 900
Db   1028 CGCACGCCAGAGGTTGCCCGCGCCGCCACCCCGTTATTGGATTTGATCAAGACACAGCG 1087

Qy    901 TTGACGCGCCCCATCCACGCGGATAACACGCCGTATGGTGTGACATTACCCACTTCAGTGCTG 960
Db   1088 TTGACGCGCCCCCATCCACGCGGATAACACGCCGTATGGTGTGACATTACCCACTTCAGTGCTG 1147

Qy    961 TTTATTGCCGGACACGATACTAATCTTGGCAAATCTCGGCGGCCACTGGAGCTCAACTGG 1020
Db   1148 TTTATCGCCGGACACGATACTAATCTCGGCAAATCTCGGCGGCCACTGGAGCTCAACTGG 1207

Qy   1021 ACGCTTCCCGGTCAGCCGGATAACACCCCGCCAGGTGGGTGAACTCGGTGTTTCAGTTGG 1080
Db   1208 ACGCTTCCCGGTCAGCCGGATAACACTAATCTCGGCAAATCTCGGCGGCCACTGGAGCTCAACTGG 1267
```

RESULT 9
US-10-601-319-7
; Sequence 7, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING THEM
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (188)...(1483)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 403
; OTHER INFORMATION: n = A,T,C or G
US-10-601-319-7

Query Match   97.4%; Score 1287.8; DB 8; Length 1901;
Best Local Similarity   99.5%;
Matches 1293; Conservative   0; Mismatches   7; Indels   0; Gaps   0;

```
Qy    1 ATGAAACGATCTTAATCCCATTTTATCTTCTTCGTAACCCGCGAAAATCTGCA   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  188 ATGAAACGGATCTTAATCCCATTTTATCTTCTTCGATTCAGTTTACGCAG   247

Qy   61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTTGGTGTGAATTGTTCAGTCGTCGTCATGGT  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  248 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTTGGTGTGAATTGTTCAGTCGTCGTCATGGT  307

Qy  121 GTGCGTGCTCCAACCAAGGCCACGGCGAAACTGATNGCAGGATGTCAACCCGGACGACGGCCA  180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  308 GTGCGGTGCTCCAACCAAGGCCACGGCGAAACTGATNGCAGGATGTCAACCCGGACGACGGCCA  367

Qy  181 ACCTGGCCGGGTAAAACTGGGTTGGCTGACACCGGCGNGGGCTAATCGGCCTATCTC  240
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  368 ACCTGGCCGGGTAAAACTGGGTTGGCTGACACCGGCGNGGGCTAATCGGCCTATCTC  427

Qy  241 GGACATTACCAACGCCAGCCGTCTGGTAGCCGACGAATTGCTGGGAAAAGGGCTGCCCG  300
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  428 GGACATTACCAACGCCAGCCGTCTGGTAGCCGACGAATTGCTGGGAAAAGGGCTGCCCG  487

Qy  301 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTGACGACGGATTTGTGAGAT  360
         |||||||||||||||||||||||||||||||||||||||||||
Db  488 CAGTCTGGTCAGGTCGCGATTATTGCTGACGACGGATTTGTGAGAT  547

Qy  361 GCCTTCGCCGGCGGCTGGCACCTGACCTGTGCAATAACCGTACATACCCAGGCAGATACG  420
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  548 GCCTTCGCCGGCGGCTGGCACCTGACCTGTGCAATAACCGTACATACCCAGGCAGATACG  607

Qy  421 TCCAGTCCCGATCCGTTATTGAATGATCTCTAAAAACTGGGCGTTTGCCAACTGGATAACGCG  480
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  608 TCCAGTCCCGATCCGTTATTGAATGATCTCTAAAAACTGGGCGTTTGCCAACTGGATAACGCG  667

Qy  481 AACGTGACTGACGGCGATCCTCAGCAGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT  540
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  668 AACGTGACTGACGGCGATCCTCAGCAGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT  727

Qy  541 CGGCAAAACGCGTTTCGCGAACTGGAAACGGGTGCTTAATTTTCCCAAATCAAACTTGTGC  600
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  728 CGGCAAAACGCGTTTCGCGAACTGGAAACGGGTGCTTAATTTTCCCAAATCAAACTTGTGC  787

Qy  601 CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCATTAACGGCAGGAACTGTTACCATCGGAACTC  660
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  788 CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCATTAACGGCAGGAACTGTTACCATCGGAACTC  847

Qy  661 AAGGTGAGCGCCGACAATGTCTCTCATTAACGGTGCGCGGTAAGCCTCGCATCAAATGTGACG  720
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  848 AAGGTGAGCGCCGACAATGTCTCATTAACGGTGCGCGGTAAGCCTCGCATCAAATGTGACG  907

Qy  721 GAGATATTTCTCCTTGCAACAAGCACAGGGAAATGCGGACCGGCGGTGGGGAAGAGATCACC  780
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  908 GAGATATTTCTCCTTGCAACAAGCACAGGGAAATGCGGACCGGCGGTGGGGAAGAGATCACC  967

Qy  781 GATTCACCAGTGGAACACCTTGCTAAGTTTGCATAAGCGCGCAATTTTATTTGGCTACAA  840
         |||||||||||||||||||||||||||||||||||||||||||
Db  968 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAAGCGCGCAATTTTATTTGGCTACAA  1027

Qy  841 CGCACGCAGGAGGTTGCCGCGCCAGCCGCGCCCGTTATTGGATTGATCATGGCAGCG  900
         ||||||||||||||||||||||||||||||||||||||||||||||
Db 1028 CGCACGCAGGAGGTTGCCGCGCCAGCCGCGCCCGTTATTGGATTGATCATGGCAGCG  1087

Qy  901 TTGACGCCCCATCCACCCGCAAAAAACAGGGCGTATGGTGACATTAACCCACTTCAGTGTACTG  960
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 1088 TTGACGCCCCATCCACCCGCAAAAAACAGGGCGTATGGTGACATTACCCACTTCAGTGCTG  1147

Qy  961 TTTATTGCCGGACACGATACTAATCTGGGCAAATCTGGGCGGCGGCGCCACTGGAGCTCAACTGG  1020
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 1148 TTTATTGCCGGACACGATACTAATCTGGGCAAATCTGGGCGGCGGCGCCACTGGAGCTCAACTGG  1207

Qy 1021 ACGCTTCCCGGGTCAGCCGAGCTGAAACAGTGGTGGTGAACGGCTGG  1080
         ||||||||||||||||||||||||||||||||||||||||||||
Db 1208 ACGGCTTCCCGGGTCAGCCGGATTAACACGCGCCCAGGTGGTGAACTGGGTGAACGGCTGG  1267

Qy 1081 CGTCGGCTAAGCGATAACAGCCAGTAGGATTCAGGTTCGGCTGGTGGTCTTCAGACTTTACAG  1140
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 1268 CGTCGGCTAAGCCGATAACAGCCAGTAGGATTCAGGTTCGGCTGGTGGTCTTCAGACTTTACAG  1327

Qy 1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGGCCGCCCCCGGGAGAGGGTGAAACTGACC  1200
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 1328 CAGATGCGTGATAAAACGCCGCTCATTAAATACGGCCGCCCCCGGGAGAGGGTGAAACTGACC  1387

Qy 1201 CTGCCAGGATGTGAAGACGGAAAATGCCAGGGCATGTGTTCGTTGCTTGTTCAGTTTTACGGCAA  1260
         ||||||||||||||||||||||||||||||||||||||||||||||
Db 1388 CTGGCCAGGATGTGAAGACGGAAAATGCCAGGGCATGTGTTCGTTGCTTGTTCAGTTTTACGGCAA  1447

Qy 1261 ATCGTGAATGAAGCACGCATACCCGGCCGTGCAGTTTGTGAGAT  1300
         |||||||||||||||||||||||||||||||
Db 1448 ATCGTGAATGAAGCACGCATACCCGGCCGTGCAGTTTGTAAT  1487
```

```
Qy  1201 CTGGGCAGGGATGTGAAGGAGGGCGAAAATGGGCAGGGCCAGGCCAGCAGGGCTTTTTACGGCAA 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1388 CTGGGCAGGGATGTGAAGAGAGGGCGAAATGCCCAGGCCAGGGCCATGTTGTCGGTGGGCGGGTTTTACGGCAA 1447

Qy  1261 ATCGTGAATGAAGCACGGCATACCGGCGGGTGCAGTTTGTGAGAT 1300
             ||
Db  1448 ATCGTGAATGAAGCACGGCATACCGGGGGGTGCAGTTTGTAAT 1487
```

RESULT 10
US-10-933-115-7

; Sequence 7, Application US/10933115
; Publication No. US20050281792A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 564462001822
; CURRENT APPLICATION NUMBER: US/10/933,115
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (188)...(1483)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 403
; OTHER INFORMATION: n = A,T,C or G
US-10-933-115-7

Query Match   97.4%; Score 1287.8; DB 10; Length 1901;
Best Local Similarity   99.5%; Pred. No. 0;
Matches 1293; Conservative   0; Mismatches   7; Indels   0; Gaps   0;

```
Qy     1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTTCTGATTCGTTAACCCGCAAATCTGCA 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTTCTGATTCGTTAACCCGCAAATCTGCA 247

Qy    61 TTCGCTCAGAGTGAGCCGGAGTGAGCTGAAAGTGTCAGTTCGTCAGTTCGTCATGGT 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248 TTCGCTCAGAGTGAGCCGGAGTGAGCTGAAAGTGTCAGTTCGTCAGTTCGTCATGGT 307

Qy   121 GTGCGTCGTCTCCAACCAAGGCCACGCAACGATCAGTCAGGAGAGTGTCACCCCAGAGGCCATGGCCA 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308 GTGGCGTGCTCCAACCAAGGCCCACGCAACTGATGCAGGAATGTCACCCCAGAGGCCATGGCCA 367

Qy   181 ACCTGGCCGCGGTAAAACTGGGTTGGGCTGACACACCGCACCGCGGGCGNGGTGGGTGAGCTAATCGGCCTATCTC 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368 ACCTGGCCGCGGTAAAACTGGGTTGGGCTGACACACCGCACCGCGGGCGGGTGGGTGAGCTAATCGGCCTATCTC 427

Qy   241 GGACATTACCAACGCCAGGCGTCTTGGTAGCCCACGACGGATTGCTTGGGGTAGAGGAAAAAAGGGCTGCCG 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 GGACATTACCAACGCCAGGCGTCTTGGTAGCCCACGACGGATTGCTTGGGGTAGAGGAAAAAAGGGCTGCCG 487

Qy   301 CAGTCTGGTCAGGTCGGGGATTATTGCTTGATGTGTCGACAGGAGAGGCGTACCCGTAAAACAGGGGAA 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488 CAGTCTGGTCAGGTCGGGGATTATTGCTTGATGTGTCGACAGGAGAGGCGTACCCGTAAAACAGGGGAA 547

Qy   361 GCCTTCGCGCCCGGGCTGGCACCTGACTGTGTGCAAATAACCGTAACTACCCAGGCGAGATACG 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548 GCCTTCGCGCCCGGGCTGGCACCTGACTGTGTGCAAATAACCGTAACTACCCAGGCGAGATACG 607

Qy   421 TCCAGTCCCGATCCGTTATTTTATTTAATCCTCTAAAAAACTGGCGTTTGCCAACTGGATAACGCG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   608 TCCAGTCCCGATCCGTTATTTTATTTAATCCTCTAAAAAACTGGCGTTTGCCAACTGGATAACGCG 667

Qy   481 AAGGTGACTGACGGCGATCCTCAGCACAGGGGCGCAGAGAGGGTCAATTGCTGACTTTACCGGGCAT 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668 AAGGTGACTGACGGCGATCCTCAGCACAGGGGTCAAATTGCTGACTTTACCGGGCAT 727

Qy   541 CGGCAAACGGCGGTTCGGGAACTGGGAACGGGTGCTTAATTTCCGGAATCAAAACTTGTGC 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   728 CGGCAAAACGGCGGTTTCGGGAACTGGGAACGGGTGCTTAATTTTCCGGAATCAAAACTTGTGC 787

Qy   601 CTTAAAACGTGAGAAAACAGGGACAGAAAGCTTGTTCATTAAACGCAGGCCATTACCATCGGGAACTC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   788 CTTAAAACGTGAGAAAACAGGGACAGAAAGCTTGTTCATTAAACGCAGGCCATTACCATCGGGAACTC 847

Qy   661 AAGGTGACGCGCGACAATGTCTCCATTAAACGGTGGCGTTAAGCCTTCGGCCATCAATGCTCACG 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   848 AAGGTGAGGCGCGACAATGTCTCCATTAAACGGTGGCGTTAAGCCTTCGGCCATCAATGCTCACG 907

Qy   721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGTGGGGAGGATCACC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   908 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGTGGGGAGGATCACC 967

Qy   781 GATTCACACCAGTGGAACAACCTTGCTTAAGTTTGCATAAACGCCAAATTTATTTGCTACAA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   968 GATTCACACCAGTGGAACAACCTTGCTTAAGTTTGCATAAACGCCAAATTTTATTTGCTACAA 1027

Qy   841 CGGCACGCCAGAGGTTGCCCGCGACCCACCCCGTTATTCGGATTTGGATCAGGCAGCG 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1028 CCGCACGCCAGAGGTTGCCCGCGACCCACCCCGTTATTCGGATTTGATCAAGACAGCG 1087

Qy   901 TTGACGCGCCCATCCACCGCAAAAACAGGGCGTATGGTGGTGACATTACCCACTTCAGTACTG 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1088 TTGACGCGCCCATCCACCGCAAAAACAGGGCGTATGGTGGTGACATTACCCACTTCAGTGCTG 1147

Qy   961 TTTATTCGCCGGACACGATACTAAATCTGCAAATCTCGCGCGGCACTGGAGCTCAACTGG 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1148 TTTATTCGCCGGACACGATACTAATTCTGGCAAATCTCGCGCGGCACTGGAGCTCAACTGG 1207

Qy  1021 ACGGCTTCCCGGTTCAGCGCCGGGATAAACACCGCCAGGCGTTGGTGTTTGAACGCGTGG 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1208 ACGGCTTCCCGGTTCAGCGCCGGGATAAACACCGCCAGGCGTTGGTGTTTGAACGGTTACAG 1267

Qy  1081 CGTCGGCGCTAAGCGATAACAGCCAGTGGGATTCAGGTTTCGCTGGTCTTCCAGTCAGGCGACTTTACAG 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1268 CGTCGGCGCTAAGCGATAACAGCCAGTGGGATTCAGGTTTCGCTGGTCTTCCAGTCAGGCGACTTTACAG 1327

Qy  1141 CAGATGCGTGATAAAAACGCGCGTGTCATTAAATACGGCGCCGCCGGAGAGGTGAAACTGACC 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1328 CAGATGCGTGATAAAAACCCGCTGTCATTAAATACGGCGCCCGGAGAGGTGAAACTGACC 1387

Qy  1201 CTGGGCAGGGATGTGAAGGAGGGCGAAATGCCGAGGCCGGCCATGTGTTTGTCGGTTTTACCGGCAA 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1388 CTGGGCAGGGATGTGAAGGAGGGCGAAATGCCGAGGCCGGCCATGTGTTTGTCGGTTTTACCGGCAA 1447

Qy  1261 ATCGTGAATGAAGCACGGCATACCGGCGGGTGCAGTTTGTGAGAT 1300
             ||
Db  1448 ATCGTGAATGAAGCACGGCATACCGGGCGGGTGCAGTTTGTAAT 1487
```

RESULT 11
US-11-056-354-3
; Sequence 3, Application US/11056354
; Publication No. US20050246780A1
; GENERAL INFORMATION:
; APPLICANT: SHORT, Jay M.
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin A.
; APPLICANT: BARTON, Nelson R.
; APPLICANT: GARRETT, James B.
; APPLICANT: O'DONOGHUE, Eileen
; APPLICANT: MATHUR, Eric J.
; TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM AND METHODS
; TITLE OF INVENTION: FOR MAKING AND USING THEM
; FILE REFERENCE: 564462001803/D1370-9C1
; CURRENT APPLICATION NUMBER: US/11,056,354
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 10/156,660
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.1
;
; SEQ ID NO 3
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli appA phytase
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (403)..(403)
; OTHER INFORMATION: n is any nucleotide
US-11-056-354-3

Query Match   97.4%;
Best Local Similarity 99.5%;
Matches 1293; Conservative 0;   Mismatches 7;   Indels 0;   Gaps 0;

Qy   1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCGCAATCTCCA   60
Db   188 ATGAAAGGGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCGCAATCTCCA 247

Qy   61 TTCGCTCAAGAGTGAGCCCGGACGCTGGAAGCGTGGCGAAAGTGTGGGTGATTCGTCGTCATGGT 120
Db   248 TTCGCTCAGAGTGAGCCCGGACTGAAGCTGGAAAAGTGTGGGTGATTGTTCAGTCTCGTCATGGT 307

Qy   121 GTGCGTGCGTCCAACCAAGGCCAGCCAACTGATGATGCAGGAGATGTCACCGGACGGCATGGGCCA 180
Db   308 GTGCGTGCGTCCAACCAAGGCCAGCCAACTGATGATGCAGGAGATGTCACCGGACGGCATGGGCCA 367

Qy   181 ACCTGGCGCGGTAAAACTGGGTTGGCTGACACCCCGCGNGNGTGGGGTGGGGGTGAGCTAATCGCCTATCTC 240
Db   368 ACCTGGCGCGGTAAAACTGGGTTGGCTGACACCCCGCGNGTGGGGTGAGCTAATCGCCTATCTC 427

Qy   241 GGACATTACCAACGCGCAGCGGTCTTGGGTAGCCGACGAGGCGGAAAAGGGGCTGCCCG 300
Db   428 GGACATTACCAACCCCCAGCGGTCTTCGGTAGCCGACGAGGCATTGCTCGGCGCGTACCCGGTAAAAACAGGGGCTGCCCG 487

Qy   301 CAGTCTGGTTCAGGTCGCGGATTATTGCTCAATGTCGACGAGCGTACCCGGTAAAAACAGGGGAA 360
Db   488 CAGTCTGGTTCAGGTCGCGGATTATTGCTCAATGTCGACGAGCGTACCCGGTAAAAACAGGGCGAA 547

Qy   361 GCCTTCGCGCCGCCGGGCTGGACCTGACTGTGCAATAACCGTACATACCCAGGCAGAGATACG 420
Db   548 GCCTTCGCGCCGCCGGGCTGGACCTGACTGTGCAATAACCGTACATACCCAGGCAGAGATACG 607

Qy   421 TCCAGTCCCGATCCGTTATTTATTAATCCTCTAAAAAACTGGCGTTTGCCAACTGGGATAACGCG 480
Db   608 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAAACTGGCGTTTGCCAACTGGGATAACGCG 667

Qy   481 AACGTGACTGAACGGATCCTTCAGCAGAGGGTCAATTGCTGGTCTTGCTGGACTTTACCGGGCAT 540
Db   668 AACGTGACTGAACGGATCCTTCAGCAGAGGGTCAATTGCTGGACTTTACCGGGCAT 727

Qy   541 CGGCAAACGCGTTTCGCGGAACTGGAACGGGTGGCTTAAATTTTCCCGCAATCAAAACTTGTGC 600
Db   728 CGGCAAACGCGTTTCGCGGAACTGGAACGGGTGGCTTAAATTTTCCCGCAATCAAAACTTGTGC 787

Qy   601 CTTAAACGTGAGAACACAGGACGGAAAGCTGTTCATTAACCGCAGCACATTACCATCGGAACTC 660
Db   788 CTTAAACGTGAGAACACAGGACGGAAAGCTGTTCATTAACCGCAGCACATTACCATCGGAACTC 847

Qy   661 AAGGTGAGCGCCGGACAAGTCTCATTAACCGGTGGGTAAGCCTCGGCATCAAGTGCTGACG 720
Db   848 AAGGTGAGCGCCGGACGACATGTCTCAGCAGGGTAAGCCTCGGCATCAATGGCTGACG 907

Qy   721 GAGATATTTCTCTTGCCAACAAGCACAGGGAAATGCGGAGCCGGGGAAGGATCACC 780
Db   908 GAGATATTTCTCCTCGCAACAAGCACAGGGAAATGCGGAGCCGGGGTGGGGAAGGATCACC 967

Qy   781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAAGCGCGAATTTTATTGTGCTACAA 840
Db   968 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAAGCGCGAATTTTATTGTGCTACAA 1027

Qy   841 CGCACGCCAGAGGTTGCCCGCCCACCCCCGTTATTGATTGATCAGTGGCAGCG 900
Db   1028 CGCACGCCAGAGGTTGCCCGCCCACCCCCGTTATTGATTGATCAAGACAGCG 1087

Qy   901 TTGACGGCCCCATCCACCGCAAAAAAACAGGGGTATGGTGGTGACATTCAGTACTG 960
Db   1088 TTGACGGCCCCATCCACCGCAAAAAAACAGGGGTATGGTGGTGACATTCAGTACTG 1147

Qy   961 TTTAAGTGCCGGACACGATACTAAATCTGGCAAATCTCGGGGCGCGCACTGAGGCTCAACTGG 1020
Db   1148 TTTATCGCCGGACACGATACTAAATCTGGCAAATCTCGGGGCGCGCACTGAGGCTCAACTGG 1207

Qy   1021 ACGCTTCCCGGTCAGCGGGATAACACGCGGCGCCAGGTGGTGGAACTGGGTGTTTGGAACGCTGG 1080
Db   1208 ACGCTTCCCGGTCAGCGGGATAACACGCGGCGCCAGGTGGTGGAACTGGGTGTTTGGAACGCTGG 1267

Qy   1081 CGTCGGCTAAGCGTGACGATAACAGCCACTGGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140
Db   1268 CGTCGGCTAAGCGTGACGATAACAGCCACTGGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1327

Qy   1141 CAGATGCGTGATAAAACGCGCGTCGTCATTAAAATACGGCCCCGAGAGTGAAAACTGACC 1200
Db   1328 CAGATGCGTGATAAAACGCGCGTCGTCATTAAAATACGGCCCCGAGAGTGAAAACTGACC 1387

Qy   1201 CTGGCAGGATGTGAAGAGCGGAAATGCGCAGGGCAAGTGGTTTTGGTTGGCAGGTTTACGCAA 1260
Db   1388 CTGGCAGGATGTGAAGAGCGGAAATGCGCAGGGCAAGTGGTTTTGGTTGGCAGGTTTACGCAA 1447

Qy   1261 ATCGTGAATGAAGCACGGCATACCCGGCGTGCAGTTTGAGAT 1300
Db   1448 ATCGTGAATGAAGCACGGCATACCCGGCGTGCAGTTTGTGAAT 1487

RESULT 12
US-10-282-122A-7167
; Sequence 7167, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl

```
;   APPLICANT: Zyskind, Judith
;   APPLICANT: Wall, Daniel
;   APPLICANT: Trawick, John
;   APPLICANT: Carr, Grant
;   APPLICANT: Yamamoto, Robert
;   APPLICANT: Forsyth, R.
;   APPLICANT: Xu, H.
;   TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;   FILE REFERENCE: ELITRA.034A
;   CURRENT APPLICATION NUMBER: US/10/282,122A
;   CURRENT FILING DATE: 2003-02-20
;   PRIOR APPLICATION NUMBER: 60/191,078
;   PRIOR FILING DATE: 2000-03-21
;   PRIOR APPLICATION NUMBER: 60/206,848
;   PRIOR FILING DATE: 2000-05-23
;   PRIOR APPLICATION NUMBER: 60/207,727
;   PRIOR FILING DATE: 2000-05-26
;   PRIOR APPLICATION NUMBER: 60/230,335
;   PRIOR FILING DATE: 2000-09-06
;   PRIOR APPLICATION NUMBER: 60/230,347
;   PRIOR FILING DATE: 2000-09-09
;   PRIOR APPLICATION NUMBER: 60/242,578
;   PRIOR FILING DATE: 2000-10-23
;   PRIOR APPLICATION NUMBER: 60/253,625
;   PRIOR FILING DATE: 2000-11-27
;   PRIOR APPLICATION NUMBER: 60/257,931
;   PRIOR FILING DATE: 2000-12-22
;   PRIOR APPLICATION NUMBER: 60/267,636
;   PRIOR FILING DATE: 2001-02-09
;   PRIOR APPLICATION NUMBER: 60/269,308
;   PRIOR FILING DATE: 2001-02-16
;   Remaining Prior Application data removed - See File Wrapper or PALM.
;   NUMBER OF SEQ ID NOS: 78614
;   SOFTWARE: Patentin version 3.1
;
;   SEQ ID NO 7167
;   LENGTH: 1299
;   TYPE: DNA
;   ORGANISM: Escherichia coli
US-10-282-122A-7167

Query Match          97.4%;  Score 1287;  DB 8;  Length 1299;
Best Local Similarity  99.5%;  Pred. No. 0;  Mismatches 0;
Matches 1290;  Conservative 0;  Mismatches 6;  Indels 0;  Gaps 0;

Qy        1 ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTCTTCTGATTCCGTTAACCCCGCAATCTGCA  60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTCTTCTGATTCCGTTAACCCCGCAATCTGCA  60

Qy       61 TTCGCTCAGAGTGAGCGCGGAGCTGAAGCTGGAAAAGTGTGGTTGAGTTCAGTCGTCATGGT  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 TTCGCTCAGAGTGAGCGCGGAGCTGAAGCTGGAAAAGTGTGGTTGAGTTCAGTCGTCATGGT  120

Qy      121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACCCATGGCCA  180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACCCATGGCCA  180

Qy      181 ACCTGGCGGGTAAAACTGGGTTGGCTCAACAACCGCGCGCGGNGGTGGGTTGAGCTTAATC  240
            ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db      181 ACCTGGCGGGTAAAACTGGGTTGGCTCAACAACCGCGCGCGGTGGGTTGAGCTTAATCCGCCTATCTC  240

Qy      241 GGACAATTACCAACGCCAGCGTCTGGTAGCCCAGCGGATTGCTCGGCGAAAAAGGGCTGCCCG  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 GGACAATTACCAACGCCAGCGTCTGGTAGCCCAGCGGATTGCTCGGCGAAAAAGGGCTGCCCG  300

Qy      301 CAGTCTGGTCAGGTCCGGATTATTGCTGAATGTCGACGAGCGTACCCGTAAAACAGGCGAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 CAGTCTGGTCAGGTCCGGATTATTGCTGAATGTCGACGAGCGTACCCGTAAAACAGGCGAA  360

Qy      361 GCCTTCGCCGCGCGGGGCTGGCACCTGACTGTGCAATAACCGTTGCATACCCAGGCAGATACG  420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 GCCTTCGCCGCGCGGGGCTGGCACCTGACTGTGCAATAACCGTTGCATACCCAGGCAGATACG  420

Qy      421 TCCAGTCCCGATCCGTTATTTAATCTCTCTAAAAAACTGGGCGTTTGCCAACTGGATAACGCG  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAAACTGGCGTTTGCCAACTGGATAACGCG  480

Qy      481 AACGTGACTGACGGCGATCCTCAGCCAGGGCAGGAGGGGTCAAATGCCGGCAT  540
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 AACGTGACTGACGGCGATCCTCAGCCAGGGCAGGAGGGGTCAAATGCCGGCAT  540

Qy      541 CGGCAAACGGCCGTTTCCGGAACTGGAACGGGGTGCTTAATTTTCCGCAATCAAACTTGTGC  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 CGGCAAACGGCCGTTTCCGGAACTGGAACGGGGTGCTTAATTTTCCGCAATCAAACTTGTGC  600

Qy      601 CTTAAACGTGAGGAAAACAGGACGAAAAGCTGTTCATTAAGCGCAGGCATTACCATCGGAACTC  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 CTTAAACGTGAGGAAAACAGGACGAAAAGCTGTTCATTAAGCGCAGGCATTACCATCGGAACTC  660

Qy      661 AAGGTGAGCGCGCGACAGATGTCTCATTAAACCGTGCGTAAGCCTCGCATCAATGCTGACG  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 AAGGTGAGCGCGCGACAGATGTCTCATTAAACCGTGCGTAAGCCTCGCATCAATGCTGACG  720

Qy      721 GAGATATTTCTCCTGCAACAACAGCACACGGGAAATGCCGGAGCCCGGGGTGGGGGAAGGATCACC  780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 GAGATATTTCTCCTGCAACAACAGCACACGGGAAATGCCGGAGCCCGGGGTGGGGGAAGGATCACC  780

Qy      781 GATTCACACCAGTGGGAACACCTTGCTAAGGTTGCATAACGGCGCGCAATTTTATTTGCTACAA  840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 GATTCACACCAGTGGGAACACCTTGCTAAGGTTGCATAACGGCGCGCAATTTTATTTGCTACAA  840

Qy      841 CGCACGCCAGAGGTTGCCCCGCAGCCGGCCACCCGTTATTAGATTTGATCAAGACACGCG  900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 CGCACGCCAGAGGTTGCCCCGCAGCCGGCCACCCGTTATTAGATTTGATCAAGACACGCG  900

Qy      901 TTGACGCCCCATCCACCGCAAAAACACGGGCGTATGGTCGTGACATTACCCACTTCAGTACTG  960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 TTGACGCCCCATCCACCGCAAAAACACGGGCGTATGGTGTGACATTACCCACTTCAGTGCTG  960

Qy      961 TTTATTCGGCGGACACGATACTAATCTCGGCAAATCTCGGCGGCACTGGAGCTCAACTGG  1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 TTTATTCGGCGGACACGATACTAATCTCGGCAAATCTCGGCGGCACTGGAGCTCAACTGG  1020

Qy     1021 ACGCTTCCCGGTCAGCGCGGATAAACACCCCGGGTGGTGAACGTGTTGAAACGCTGG  1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 ACGCTTCCCGGTCAGCGCGGATAAACACCCCGGGTGGTGAACGTGTTGAAACGCTGG  1080

Qy     1081 CGTCGGCTAAGGCGTAACAGCCAGTGGAATTCAGGTTTCGGCTGGTCTTCCAGACTTTACAG  1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 CGTCGGCTAAGGCGTAACAGCCAGTGGAATTCAGGTTTCGGCTGGTCTTCCAGACTTTACAG  1140

Qy     1141 CAGATGGCTGATAAAAACGCCCGCTGTCATTAAATACGGCGCCCGGAGAGGTGAAACTGACC  1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 CAGATGGCTGATAAAAACGCCCGCTGTCATTAAATACGGCGCCCGGAGAGGTGAAACTGACC  1200

Qy     1201 CTGGCAGGATGTGAAGAGCGAAATGCCGCAGGGCCATGTNGTTCCTTGGCAGGTTTTACCCAA  1260
            |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db     1201 CTGGCAGGATGTGAAGAGCGAAATGCCGCAGGGCCATGTCGTTCCTTGGCAGGTTTTACCCAA  1260

Qy     1261 ATCGTGAATGAAGCACCCATACCGGGCCTGCAGTTTG  1296
            |||||||||||||||||||||||||||||||||||||
Db     1261 ATCGTGAATGAAGCACCCATACCGGGCCTGCAGTTTGCACGTTTG  1296
```

```
RESULT 13
US-09-866-379-5
;   Sequence 5, Application US/09866379
;   Patent No. US2002013675A1
;   GENERAL INFORMATION:
;   APPLICANT: DIVERSA CORPORATION
;   APPLICANT: SHORT, Jay
;   APPLICANT: KRETZ, Keith
;   APPLICANT: GRAY, Kevin
;   APPLICANT: BARTON, Nelson
;   APPLICANT: GARRETT, James
;   APPLICANT: O'DONOHUE, Eileen
;   TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
```

; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR APPLICATION FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; SOFTWARE: Patentin version 3.1
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO 5
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide

US-09-866-379-5

Query Match   97.3%;   Score 1286.2;   DB 3;   Length 1901;
Best Local Similarity   99.4%;   Pred. No. 0;   Mismatches   8;   Indels   0;   Gaps   0;
Matches 1292;   Conservative   0;

```
Qy      1 ATGAAAGGCGATCTTAATCCCATTTTATCTCTTCTGATTCCGTTAACCCTGCAATCTGCA     60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188 ATGAAAGGGATCTTAATCCCATTTTATCTCTTCTGATTCCGTTAACCCTGCAATCTGCA    247

Qy     61 TTCGCTCAGAGTGAGCCCGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT    120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248 TTCGCTCAGAGTGAGCCCGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT    307

Qy    121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAAGATGTCACCCCAGACGCATGGCCA    180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAAGATGTCACCCCAGACGCATGGCCA    367

Qy    181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCCGCGNGTGGTTGAGCTAATCGCCTATCTC    240
            |||||||||||||||||||||||| |||||||||  ||||| ||||||||||||||||
Db    368 ACCTGGCCGGTAAAACTGGGTTGACTGACACCCGCGNGTGGTTGAGCTAATCGCCTATCTC    427

Qy    241 GGACATTACCAACGCCAGCGGTCTGGGTAGCCGACGGGATTGGCTGGCGAAAAAGGGCTGCCCG    300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428 GGACATTACCAACGCCAGCGGTCTGGGTAGCCGACGGGATTGCTGGCGAAAAACAGGCCGA    487

Qy    301 CAGTCTGGTTCAGGTCGCGGCGGATTATTGCTGAATGTCGATCGATGAGGCCGTACCCGGTAAAACAGGCCGAA    360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488 CAGTCTGGTTCAGGTCGCGGATTATTGCTGAATGTCGATCGATGAGGCCGTACCCGGTAAAACAGGCCGAA    547

Qy    361 GCCTTCGCCGCGCCCGGGCTGCCACTGTGACTGTTGCAATAACCGTACTACTACCAGGCCAGATACG    420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548 GCCTTCGCCGCGCCCGGGCTGGCACTGTGCAATAACCGTACTACTACCAGGCCAGATACG    607

Qy    421 TCCAGTCCGATCCGGTTATTTAATTCTCTAAAAACTGGATCCGGATAACGGG    480
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608 TCCAGTCCGATCCGGTTATTTAATTCCTAAAAACTGGCGTTGCCAACTGGATAACGCTGG    667

Qy    481 AACGTGACTGACGCGATCCTCAGCAGGGCGCAGAGAGGGTCAATTGCTTGACTTTACCGGGGCAT    540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    668 AACGTGACTGACGCGATCCTCAGCAGGGCGCAGAGGGGTCAATTGCTTGACTTTACCGGGGCAT    727

Qy    541 CGGCAAACGGCCGTTTCGCGAACTGGAACGGGTCTTAATTTTCCGCAATCAAAACTTGTTGC    600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    728 CGGCAAACGGCCGTTTCGCGAACTGGAACGGGTCTTAATTTTCCGCAATCAAAACTTGTTGC    787

Qy    601 CTTAAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCACAGGAGCATTACCAATCGGGAACTC    660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    788 CTTAAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCACAGGAGCATTACCAATCGGGAACTC    847
```

RESULT 14
US-09-866-379-9
; Sequence 9, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1

```
Qy    661 AAGGTGAGCGGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGGCGTAAGCTGACG    720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    848 AAGGTGAGCGGCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGGCGTAAGCTGACG    907

Qy    721 GAGATATTTCTTCCTGCAACAAGCACAGGAGAATGCCGGAGCAGGGGATCACC    780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908 GAGATATTTCTTCCTGCAACAAGCACAGGAGAATGCCGGAGCAGGGGAGGGATCACC    967

Qy    781 GATTCACACCAGTGAACACCTTGCTAAGTTTGCATTAAACCGGCAATTTATTTTGCTACAA    840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968 GATTCACACCAGTGAACACCTTGCTAAGTTTGCATTAAACCGGCAATTTATTTTGCTACAA    1027

Qy    841 CGCACGCCAGGAGGTTGCCCCGGCCGCGCACCCCGGTTATTAGGATTGATCAAGACAGCG    900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028 CGCACGCCAGGAGGTTGCCCCGGCCGCGCACCCCGGTTATTAGGATTGATCAAGACAGCG    1087

Qy    901 TTGACGCCCCATCCACCGCAAAAAACAGGCCGTATGGTGACATTACCCACTTCAGTACTG    960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 TTGACGCCCCATCCACCGCAAAAAACAGGCCGTATGGTGACATTACCCACTTCAGTACTG    1147

Qy    961 TTTATTGCCGGACACGATACTACTAATCTTGGCAAATCTCGGCGGCTCAACTGG    1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148 TTTATCGCCGGACACGATACTACTAATCTTGGCAAATCTCGGCGGCTCAACTGG    1207

Qy   1021 ACGCTTCCCGTGCTCAGCCGGATAACACGCCCGCCAGTGGTGGTGAACTGGTGTTGAACGCTGG    1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208 ACGCTTCCCGTGCTCAGCCGGATAACACGCCCGCCAGTGGTGGTGAACTGGTGTTGAACGCTGG    1267

Qy   1081 CGTCGGCGCTAAGCGGATAACAGCCACTGGATTCAGGTTTCGGTCTGGTGCTTCTTCAGACTTTACAG    1140
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268 CGTCGGCGCTAAGCGGATAACAGCCACTGGATTCAGGTTTCGGTCTGGTGCTTCTTCAGACTTTACAG    1327

Qy   1141 CAGATGCGTGGATAAAACGCGCTGTCATTAAAATAACGGCCGCCGGGAGAGGGTGAAAACTGACC    1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328 CAGATGCGTGGATAAAACGCGCTGTCATTAAAATAACGGCCGCCGGGAGAGGGTGAAAACTGACC    1387

Qy   1201 CTGGCAGGATGTGAAGAGCGAAATGCGAGGGCATGTGTTCGTTGGCAGGTTTACGCAA    1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388 CTGGCAGGATGTGAAGAGCGAAATGCGAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA    1447

Qy   1261 ATCCTGAATGAAGCACGCATTACCCGCGTCGGTCAGTTTGAGAT    1300
            ||||||||||||||||||||||||||||||||||||||||||||
Db   1448 ATCCGTGAATGAAGCACGCATTACCCGCGTGCAGTTTGTGAAT    1487
```

```
;
; SEQ ID NO 9
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-9

Query Match    97.3%;  Score 1286.2;  DB 3;  Length 1901;
Best Local Similarity  99.4%;  Pred. No. 0;
Matches 1292;  Conservative 0;  Mismatches 8;  Indels 0;  Gaps 0;

Qy      1 ATGAAAGCGATCTAATCCATTTTATCTTTCTTCTGATTCGTTAACCCGCCAATCTGCA   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188 ATGAAAGCGATCTTAATCCATTTTATCTTTCTTCTGATTCGTTAACCCGCCAATCTGCA  247

Qy     61 TTCGCTCAGAGTGAGCCGGAGCTGAACGTGGAAAGTGTCAGTCGTCGTCAGTGGT  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248 TTCGCTCAGAGTGAGCCGGAGCTGAACGTGGAAAGTGTCAGTCGTCGTCATGGT  307

Qy    121 GTGCGTGCTCCAACCAAGGCCACGCAACCAACTGATGAGTCACCCCAGACGCCATGGCCA  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGAGTCACCCCAGACGCCATGGCCA  367

Qy    181 ACCTGGCCGGTAAAACTGGGGTTGGCCTGACACCCCGNGCGTGACACTCGCCTATCTC  240
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    368 ACCTGGCCGGTAAAACTGGGGTTGGCCTGACACCCCGNGGGTGACACTCGCCTATCTC  427

Qy    241 GGACATTACCAAACGCCAGCGCGTCTGGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428 GGACATTACCAAACGCCAGCGTCTGGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG  487

Qy    301 CAGTCTGGTCAGGTCGGATTATTGCTGATGTCGACGAGCGTACCCCGTAAAAACAGGGGAA  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488 CAGTCTGGTCAGGTCGGATTATTGCTGATGTCGACGAGCGTACCCCGTAAAAACAGGCGAA  547

Qy    361 GCCTTCGCGCGCCGGCTGGCACCTGACTGTGTGCAATAACCGTACATACCAGGCAGATACG  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548 GCCTTCGCGCGCCGGCTGGCACCTGACTGTGTGCAATAACCGTACATACCAGGCAGATACG  607

Qy    421 TCCAGTCCCGATCCGGTTATTTAATCTCTAAAAAACTGGATGATAACGCG  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608 TCCAGTCCCGATCCGGTTATTTAATCTCTAAAAAACTGGATGATAACGCG  667

Qy    481 AACGTGACTGACGCGGATCCTCAGCAGCGGGCCAGAGGGTCAATTGCTGACTTTACCGGGCAT  540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    668 AACGTGACTGACGCGGATCCTCAGCAGCGGGCCAGAGGGTCAATTGCTGACTTTACCGGGCAT  727

Qy    541 CGGCAAACGGCGTTTCGCGGAACTGGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC  600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    728 CGGCAAACGGCGTTTCGCGGAACTGGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC  787

Qy    601 CTTAAACGTGAGAAAACAGGAGCAGAAAAGCTGTGTCATTAACGCACGCATTACCATCGGAACTC  660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    788 CTTAAACGTGAGAAAACAGGAGCAGAAAAGCTGTGTCATTAACGCACGCATTACCATCGGAACTC  847

Qy    661 AAGGTGAGCGCGACAAATGTCTCATTAACCGGTGCGGTAAGCCTCGCAATCAATGCTGACG  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    848 AAGGTGAGCGCGACAAATGTCTCATTAACCGGTGCGGTAAGCCTCGCAATCAATGCTGACG  907

Qy    721 GAGATATTTCTCTGCAACAAGCACAGGGAATGCCGGAGCCGGGTGGGGAGGGATCACC  780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908 GAGATATTTCTCTGCAACAAGCACAGGGAATGCCGGAGCCGGGTGGGGAAGGGATCACC  967

Qy    781 GATTCACACCAGTGGGAACACCTTGCTAAGTTTGCATAACGCCCAATTTTATTTGCTACACA  840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968 GATTCACACCAGTGGGAACACCTTGCTAAGTTTGCATAACGCCCAATTTTGTTTGCTACAA  1027

Qy    841 CGCACCGCCAGAGGTTGCCCGCGCCGCCACCCCGTTATTTAGATTTGATCATGGGCAGCG  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028 CGCACCGCCAGAGGTTGCCCGCGCCGCCACCCCGTTATTTAGATTTGATCATGGGCAGCG  1087

Qy    901 TTGACGCGCCCATCCACCGGCACCGCGCAAAAACAGGCGTATTGGGTGTGACAATTACCCACTTCAGTACTG  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 TTGACGCGCCCATCCACCGGCACCGCGCAAAAACAGGCGTATGGGTGTGACAATTACCCACTTCAGTACTG  1147

Qy    961 TTTATTCGTCGGGACACGATACTAATCTGGCAAAATCTCGGCGGCGGCAGCTCGGCAGCTCAACTGG  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148 TTTATTCGTCGGGACACGATACTAATCTGGCAAAATCTCGGCGGCGGCAGCTCGGCAGCTCAACTGG  1207

Qy   1021 ACGCTTCCCGGTCGAGCCGGGATAACACACGCCGCCAGGTGGTGAACTGGTGGTTTGAACGTGG  1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208 ACGCTTCCCGGTCGAGCCGGGATAACACGCCGCCAGGTGGTGAACTGGTGGTTTGAACGTGG  1267

Qy   1081 CGTCGCGCTGAAGCCGGATAACAGCACGTGGAATTCAGGTTTGCGTTGCTTCTTCCAGAGACTTTACAG  1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268 CGTCGCGCTGAAGCCGGATAACAGCACGTGGAATTCAGGTTTGCGTTGCTTCTTCCAGAGACTTTACAG  1327

Qy   1141 CAGATGCGTGATAAAACCCGCGTCTGTCATTAAATAACGCGGCCCCGGAGAGGAGGTGAAAACTGACC  1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328 CAGATGCGTGATAAAACCCGCGTCTGTCATTAAATAACGCGGCCCCGGAGAGGAGGTGAAAACTGACC  1387

Qy   1201 CTGGCAGGATGTGAAAGAGCGAAATGCGCAGGGCCATGTGGTTTCGTTGGCAGGTTTTACGGCAA  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388 CTGGCAGGATGTGAAAGAGCGAAATGCGCAGGGCCATGTGGTTTCGTTGGCAGGTTTTACGGCAA  1447

Qy   1261 ATCGTGAATGAAGCACACCGCATACCGGGCGTGCGCAGTTTGAGAT  1300
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   1448 ATCGTGAATGAAGCACACCGCATACCGGGCGTGCGCAGTTTGTAAT  1487
```

RESULT 15
US-10-601-319-5
; Sequence 5, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING THEM
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
;
; SEQ ID NO 5
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 403
; OTHER INFORMATION: n = A,T,C or G
US-10-601-319-5

Query Match   97.2%; Score 1284.6; DB 8; Length 1901;
Best Local Similarity 99.3%; Pred. No. 0;

```
Matches 1291;  Conservative  0;  Mismatches  9;  Indels  0;  Gaps  0;

Qy     1 ATGAAAGCCGATCTTAATCCCATTTTTATCTTCTTCTGATTCGTTAACCCGCGCAATCTGCA    60
         ||||||| ||||||||||||||||||| ||||| || | ||||||||||||||||||||||||
Db   188 ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTCCGTTAACCCGCGCAATCTGCA       247

Qy    61 TTCGCTCAGAGTGAGCCCGGAGCTGAAGCTGGAAAGTGTGGTGATTCGTTCAGTCGTCATGGT   120
Db   248 TTCGCTCAGAGTGAGCCCGGAGCTGAAGCTGGAAAGTGTGGTGATTCAGTCGTCATGGT     307

Qy   121 GTCGGTGCTCCAACCAAGGCCACCCAACTGAATGCAGAAGTCACCCCAGACGCATGGCCA    180
Db   308 GTCGGTGCTCCAACCAAGGCCACCCAACTGAATGCAGAAGTCACCCCAGACGCATGGCCA    367

Qy   181 ACCTGGCCCGGTAAAACTGGGTTGGCTGACACCGCGCCNGACGCTAATCGCCTATCTC    240
Db   368 ACCTGGCCCGGTAAAACTGGGTGAGCTGACACCGCGCCNGGGTGAGCTAATCGCCTATCTC   427

Qy   241 GGACATTACCACGCCACGCGTCTTGGTAGCCGACGCGGATTGCTGCGGGAAAAAGGGCTGCCCG   300
Db   428 GGACATTACCAACGCCAGGGTCTTGGTAGCCGACGCGGATTGCTGCGGCGGAAAAAGGGCTGCCCG  487

Qy   301 CAGTCTGGTCAAGGTCGCCGGATTATTGCTGCAATGGTCGGACGAGCGTACCCCGTAAAACAGGCGAA   360
Db   488 CAGTCTGGTCAAGGTCGGCGGATTATTGCTGCAATGGTCGGACGAGCGTACCCCGTAAAACAGGCGAA   547

Qy   361 GCCTTCGCGCGCCGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG    420
Db   548 GCCTTCGCGCGCCGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG    607

Qy   421 TCCAGTCCCGATCCGGTTATTTAATCCTCTAAAAACTGGCGTTAACCGCGG    480
Db   608 TCCAGTCCCGATCCGGTTATTTAATCCTCTAAAAACTGGCGTTAACCGCGG    667

Qy   481 AAGGTGACTGACCGGATCCTCAGCCAGGGGCGAGGAGGTCAATTGCTGACTTTACCGGGGCAT    540
Db   668 AACGTGACTGACCGGATCCTCAGCGACAGGGGGTCAATTGCTGACTTTACCGGGGCAT    727

Qy   541 CGGCAAACGGCGTTTCGCGAACTGGAACGGGGTGCTTAATTTTCCGCAATCAAAACTTGTGC    600
Db   728 CGGCAAACGGCGTTTCGCGAACTGGAACGGGGTCTTAATTTTCCGCAATCAAAACTTGTGC    787

Qy   601 CTTAAAGGTGACAAAACAGGACAGAAAGCTGTTCATTAACGGCAGCCATTACCATCGGGAACTC    660
Db   788 CTTAAAGGTGAGAAAACAGGACAGAAAGCTGTTCATTAACGGCAGCCATTACCATCGGGAACTC    847

Qy   661 AAGGTGACGCGCCGACAAATGTCTCAATTAACCGGTGCGGTAAGCCTTCGCATCAATGCTGACG    720
Db   848 AAGGTGAGCGCCGGCCGACAAATGTCTCAATTAACCGGTGCGGTAAGCCTTCGCATCAATGCTGACG    907

Qy   721 GAGATATTTCTCCTGCAACAAGCACAGCGGAATGCCCGGAGCCGGGTGGGGAAGGATCACC    780
Db   908 GAGATATTTCTCCTGCAACAAGCACAGCGGATGCCCGGAGCCGGGGTGGGGAAGGATCACC    967

Qy   781 GATTCACCAGTGGACACCAGTCCTGCTAAGTTTGCATAACGCGCAAATTTATTTGCTACAA    840
Db   968 GATTCACACCAGTGGACACCCTTGCTAAGTTTCCATAACGCGCAATTTTATTTGCTACAA   1027

Qy   841 CGCACGCCAGAGGTTGCCCGCAGCCGGCGCCACCCCGTTATTGGATTTGATTCATGGCAGCG    900
Db  1028 CGCACGCCAGAGGTTGCCCGCAGCCGGCGCCACCCCGTTATTGGATTTGATTCATGGCAGCG   1087

Qy   901 TTGACGCCCCATCCACCCGCAAAAACAGCGTATGGGTGTTGACNTTACCCACTTCAGTACTG    960
Db  1088 TTGACGCCCCATCCACCCGCAAAAACAGCGTATGGGTGTTGACATTACCCACTTCAGTGCTG   1147

Qy   961 TTTATTCGGCGGACCACGATACTAATCTCGGCGGCCCACTGGAGCTCAACTGG   1020
Db  1148 TTTATTCGGCGGACCACGATACTAATCTCGGCGGCCCACTGGAGCTCAACTGG   1207

Qy  1021 ACGCTTCCGGTCAGCGGGGATAAACACGCGCCAGGTGGTGGTGAACTGGTGTTTGAAACGCTGG   1080
Db  1208 ACGCTTCCGGTCAGCGGGGATAAACACGCGCCAGGTGGTGGTGAACTGGTGTTTGAAACGCTGG   1267

Qy  1081 CGTCGGCTAAGCGGATAAACAGCCAGTGGGATTCAGGTTTCGGTTGGTGGTCTTTCCAGACTTTACAG   1140
Db  1268 CGTCGGCTAAGCGGATAAACAGCCAGTGGGATTCAGGTTTCGGTTGGTGGTCTTTCCAGACTTTACAG   1327

Qy  1141 CAGATGCGTGATATAAAACGCGCGCCGCTGTCATTAAAATACGGTCGCCGCCCGGAGGTTGAAACTGACC   1200
Db  1328 CAGATGCGTGATATAAAACGCGCGCCGCTGTCATTAAAATACGGTCGCCGCCCGGAGGTTGAAACTGACC   1387

Qy  1201 CTGGCGAGGATGTGAAGAGCGCAAAATCGGCAGGGGCAGTGTGGTTCGGTTGCCGGCAA   1260
Db  1388 CTGGCGAGGATGTGAAGAGCGCAAAATCGCGCAGGGCARGTGTCGTTGGCAGGGTTTTTACGGCAA   1447

Qy  1261 ATCGTGAAATGAAAGCACGCAATACCGGCGCGTGCAAGTTTGAGAT   1300
Db  1448 ATCGTGAAATGAAGAGCACGCCATACCGGCGCGTGCAAGTTTGTGTAAT   1487
```

Search completed: June 14, 2006, 12:41:35
Job time : 1618.16 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.9

Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2006, 10:04:50 ; Search time 263.494 Seconds
          (without alignments)
          9394.823 Million cell updates/sec

Title:	US-10-601-319-1
Perfect score:	1322
Sequence:	1 atgaaagcgatcttaatccc..........ataccaccatccactaa 1323
Scoring table:	IDENTITY_NUC
	Gapop 10.0 , Gapext 1.0
Searched:	1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
     Maximum Match 100%
     Listing first 45 summaries

Database : Issued_Patents NA:*
 1: /EMC_Celerra_SIDS3/ptodata/2/ina/1_COMB.seq:*
 2: /EMC_Celerra_SIDS3/ptodata/2/ina/5_COMB.seq:*
 3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A_COMB.seq:*
 4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B_COMB.seq:*
 5: /EMC_Celerra_SIDS3/ptodata/2/ina/7_COMB.seq:*
 6: /EMC_Celerra_SIDS3/ptodata/2/ina/H_COMB.seq:*
 7: /EMC_Celerra_SIDS3/ptodata/2/ina/PCTUS_COMB.seq:*
 8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP_COMB.seq:*
 9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE_COMB.seq:*
 10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1322	100.0	1323	3	US-09-259-214-1	Sequence 1, Appli
2	1322	100.0	1323	3	US-09-318-528-1	Sequence 1, Appli
3	1322	100.0	1323	3	US-09-291-931-1	Sequence 1, Appli
4	1322	100.0	1323	3	US-09-580-515-1	Sequence 1, Appli
5	1322	100.0	1323	3	US-09-866-379-1	Sequence 1, Appli
6	1287.8	97.4	1901	3	US-09-866-379-7	Sequence 7, Appl1
7	1287.8	97.4	1901	3	US-09-866-379-5	Sequence 5, Appl1
8	1286.2	97.3	1901	3	US-09-866-379-9	Sequence 9, Appl1
9	1286.2	97.3	1901	3	US-09-866-379-6	Sequence 6, Appl1
10	1283	97.0	1901	3	US-09-540-1498-9	Sequence 9, Appl1
11	1263.8	95.6	1489	4	US-10-266-041A-9	Sequence 9, Appl1
12	1263	95.5	1486	3	US-09-715-477-2	Sequence 2, Appl1
13	1251.8	94.7	1486	3	US-09-715-477-4	Sequence 4, Appl1
14	1210	91.5	1272	2	US-08-910-798-1	Sequence 1, Appl1
15	63.6	4.8	1266	3	US-09-489-039A-341	Sequence 341, App
16	39.8	3.0	2868	3	US-09-710-794-4	Sequence 4, Appl1
17	36.4	2.8	480	3	US-09-252-991A-15917	Sequence 15917, A
18	36.4	2.8	762	3	US-09-252-991A-15888	Sequence 15888, A
19	36.4	2.8	885	3	US-09-252-991A-15774	Sequence 15774, A
20	36.4	2.8	1068	3	US-09-252-991A-15804	Sequence 15804, A
21	34.6	2.6	2232	3	US-09-902-540-2830	Sequence 2830, Ap
22	34.6	2.6	14570	3	US-09-902-540-1012	Sequence 1012, Ap
23	34.4	2.6	3076	3	US-09-710-794-1	Sequence 1, Appli
24	34.4	2.6	44403765	3	US-09-103-840A-2	Sequence 2, Appli
25	34.4	2.6	4411529	3	US-09-103-840A-1	Sequence 1, Appli
26	34	2.6	998	3	US-07-885-089B-5	Sequence 5, Appli
27	34	2.6	1044	3	US-09-252-991A-8907	Sequence 8907, Ap
28	34	2.6	1068	3	US-09-710-794-3	Sequence 3, Appli
29	34	2.6	2313	3	US-09-252-991A-8696	Sequence 8696, Ap
30	34	2.6	2706	3	US-09-252-991A-9082	Sequence 9082, Ap
31	33.8	2.6	12257	3	US-09-949-016-16384	Sequence 16384, A
32	33.6	2.5	7515	3	US-09-949-016-583	Sequence 583, App
33	33.6	2.5	8976	3	US-09-621-976-8976	Sequence 8976, Ap
34	33.4	2.5	399	3	US-09-621-976-15639	Sequence 15639, A
35	33.2	2.5	1584	3	US-09-489-039A-497	Sequence 497, App
36	33.2	2.5	505	3	US-09-621-976-15639	Sequence 15639, A
37	33	2.5	505	3	US-09-252-991A-116	Sequence 116, App
38	33	2.5	2511	3	US-08-680-326-116	Sequence 116, App
39	33	2.5	35100	3	US-08-757-669A-17	Sequence 17, Appl
40	33	2.5	35100	3	US-08-770-379-17	Sequence 17, Appl
41	33	2.5	69701	3	US-09-949-016-14187	Sequence 14187, A
42	33	2.5	73308	3	US-09-949-016-16326	Sequence 16326, A
43	33	2.5	84425	3	US-09-949-016-17402	Sequence 17402, A
44	32.8	2.5	1362	3	US-09-489-039A-2259	Sequence 2259, Ap
45	32.8	2.5	1389	3	US-09-489-039A-2538	Sequence 2538, Ap

ALIGNMENTS

RESULT 1
US-09-259-214-1
; Sequence 1, Application US/09259214A
; Patent No. 6110719
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: DIVER1370-1
; CURRENT APPLICATION NUMBER: US/09/259,214A
; CURRENT FILING DATE: 1999-03-01
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; NAME/KEY: misc_feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-259-214-1

Query Match 100.0%; Score 1322; DB 3; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGAAAGCGATCTTAATCCCATATTTTATCTTCTTGATTCCGTTAACCCCGCGAATCTGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAAGCGATCTTAATCCCATATTTTATCTTCTTGATTCCGTTAACCCCGCGAATCTGCA 60

Qy         61 TTCGCTCAGAGTGAGGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTTCAGTCGTCATGGT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGCTCAGAGTGAGGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTTCAGTCGTCATGGT 120

Qy        121 GTGCGTGCTTCCAACCCAGGGCCAACCAACTTGATTCAGGATGTTCACCCCCAGACGGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTGCGTGCTTCCAACCCAGGGCCACCAACTGATCAGGATGTTCACCCCCAGACGGCATGGGCA 180

Qy        181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCCGCGGNGGTGGTGAGCTAATCGGCCTATCTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCCGCGNGGTGGTGAGCTAATCGGCCTATCTC 240
```

```
RESULT 2
US-09-318-528-1
; Sequence 1, Application US/09318528
; Patent No. 6183740
; GENERAL INFORMATION:
;  APPLICANT: Kretz, Keith
;  TITLE OF INVENTION: NOVEL PHYTASE
;  FILE REFERENCE: 09010/029003
;  CURRENT APPLICATION NUMBER: US/09/318,528
;  CURRENT FILING DATE: 1999-05-25
;  EARLIER APPLICATION NUMBER: 09/291,931
;  EARLIER FILING DATE: 1999-04-13
;  EARLIER APPLICATION NUMBER: 08/910,798
;  EARLIER FILING DATE: 1997-08-13
;  EARLIER APPLICATION NUMBER: 09/259,214
;  EARLIER FILING DATE: 1999-03-01
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;  LENGTH: 1323
;  TYPE: DNA
;  ORGANISM: Escherichia coli
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (1)..(1320)
;  NAME/KEY: misc_feature
;  LOCATION: (1)..(1323)
;  OTHER INFORMATION: n = A,T,C or G
US-09-318-528-1

Query Match       100.0%;  Score 1322;  DB 3;  Length 1323;
Best Local Similarity   100.0%;
Matches 1323;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
Qy    1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCGTTAACCCGTAACGTCGCA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCGTTAACCCGTAACGTCGCA 60

Qy   61 TTCGCTCAGAGTGAGCCCGGAGCTGAACGTGGAAAAGTGTGGGTGTTGTCATGGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCCGGAGCTGAACGTGGAAAAGTGTGGGTGTTGTCATGGT 120

Qy  121 GTGCGGTGCTCCAACCAAGGGCCAACGCCAACTGATGCAGCAGAGACCGCATGGCCA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTGCGGTGCTCCAACCAAGGGCCAACGCCAACTGATGCAGCAGAGACCGCATGGCCA 180

Qy  181 ACCTGGCCGGTAAAACTCGGGTTGGCTGACACCCCGNGGTGGTGAGCTAATCGGCCA 240
          |||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db  181 ACCTGGCCGGTAAAACTCGGGTTGGCTGACACCCCGCGGTGGTGAGCTAATCGGCCA 240

Qy  241 GGACATTACCAACGCCAGGCGTCTGGGTAGCCGACGGATTGCTGGCGGAAAAAGGGCTGCCCG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GGACATTACCAACGCCAGGCGTCTGGGTAGCCGACGGATTGCTGGCGGAAAAAGGGCTGCCCG 300

Qy  301 CAGTCTGGTCAGGTCGGCGGATTATTGCTGATGTCGACGGAGCGGAAACAGGCGGAA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGTCTGGTCAGGTCGGCGGATTATTGCTGATGTCGACGGAGCGGAAACAGGCGGAA 360

Qy  361 GCCTTCGCCGCGCCGGGCTTGGCACCTGACTGTGCAATAACCGTACCCAGGCGAGATACG 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GCCTTCGCCGCGCCGGGCTTGGCACCTGACTGTGCAATAACCGTACCCAGGCGAGATACG 420

Qy  421 TCCAGTCCCGATCCGTTATTTTAATCTTCTAAAAAACTGGATAACGCGCAT 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TCCAGTCCCGATCCGTTATTTTAATCTTCTAAAAAACTGGATAACGCGCAT 480

Qy  481 AACGTGACTGACGCGGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 AACGTGACTGACGCGGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 540
```

... (additional alignment blocks continue, Qy/Db sequence pairs from position 541 through 1323) ...

```
Qy 1321 TAA 1323
          |||
Db 1321 TAA 1323
```

```
SOFTWARE: FastSEQ for Windows Version 4.0
;
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; NAME/KEY: misc_feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-291-931-1

Query Match    100.0%;  Score 1322;  DB 3;  Length 1323;
Best Local Similarity   100.0%;  Pred. No. 0;  Mismatches 0;  Indels 0;  Gaps 0;
Matches 1323;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

RESULT 3

US-09-291-931-1

; Sequence 1, Application US/09291931A
; Patent No. 6190897
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/291,931A
; CURRENT FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
;

RESULT 4

US-09-580-515-1

; Sequence 1, Application US/0958015
; Patent No. 6720014
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/580,515
; CURRENT FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; NAME/KEY: misc_feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G

US-09-580-515-1

Query Match   100.0%;   Score 1322;   DB 3;   Length 1323;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1323;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

[DNA sequence alignment — Qy (Query) vs Db (Database), positions 1–1323]

RESULT 5
US-09-866-379-1
; Sequence 1, Application US/09866379
; Patent No. 6855365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n is any nucleotide
; NAME/KEY: CDS
; LOCATION: (1)..(1323)
; OTHER INFORMATION:
US-09-866-379-1

Query Match   100.0%;   Score 1322;   DB 3;   Length 1323;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1323;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy      1  ATGAAAACGGATCTTAATCCATTTTTATCTCTCTGATTCGTTCCGTTACCCCGCAATCTGCA  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  ATGAAAACGGATCTTAATCCATTTTTATCTCTCTGATTCGTTCCGTTACCCCGCAATCTGCA  60

Qy     61  TTCGCTTCAGAGTTGAGCCGGAGCTGAAGCTGTGGTGTAGTTGTCAGTTCGGTCGGT      120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  TTCGCTTCAGAGTTGAGCCGGAGCTGAAGCTGTGGTGTAGTTGTCAGTTCGGTCGGT     120

Qy    121  GTGGCTGGTCCAACCAAGGCCACCGCAACTGATGCAGGAGTGTCACCCGACGAC         180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  GTGGCTGGTCCAACCAAGGCCACCGCAACTGATGCAGGAGTGTCACCCGACGAC        180

Qy    181  ACCTTGGCCGGTAAAAACTGAGGTTGGCTGACACCGGCCTAATCGCCTATCTC        240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  ACCTTGGCCGGTAAAAACTGAGGTTGGCTGACACCGGCCTAATCGCCTATCTC       240

Qy    241  GGACATTACCAAACGCCAGCGCTCTCGTGTAGCGGGAATTGCTGCGGAAAGGCTG       300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  GGACATTACCAAACGCCAGCGCTCTCGTGTAGCGGGAATTGCTGCGGAAAGGCTG      300

Qy    301  CAGTCTGGTCAGGTCGGCGGATTATTGCTGAGTCGACGCGGTAAAAACAGGCTG        360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  CAGTCTGGTCAGGTCGGCGGATTATTGCTGAGTCGACGCGGTAAAAACAGGCTG       360

Qy    361  GCCTTCGCCGCGGGCTGGCCACCTGACTGCAATAACCGGAGGCAGATACG          420
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  GCCTTCGCCGCGGGCTGGCCACCTGACTGCAATAACCGGAGGCAGATACG         420

Qy    421  TCCAGTTCCCGGATCCGTTATTTAATCCTTCTAAAAACTGGGATAACGCG          480
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  TCCAGTTCCCGGATCCGTTATTTAATCCTTCTAAAAACTGGGATAACGCG         480

Qy    481  AACGTGACTGACGCGGATCCTCAGCGAGGGCAGGAGGGTCAATGCTGAG          540
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  AACGTGACTGACGCGGATCCTCAGCGAGGGCAGGAGGGTCAATGCTGAG         540

Qy    541  CGGCAAAACCGGCGTTTCGGCGAAACTGGACACCGGGTGCTTAATTTCC          600
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  CGGCAAAACCGGCGTTTCGGCGAAACTGGACACCGGGTGCTTAATTTCC         600

Qy    601  CTTAAAACGTTGAGGAACACAGAGACGAAAAGCTCTTCATTAACGGGA          660
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  CTTAAAACGTTGAGGAACACAGAGACGAAAAGCTCTTCATTAACGGGA         660

Qy    661  AAGGTGAGCGCCGACGACAATGTCTCATTAACGGTGGGGTAAGCCTTCG         720
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  AAGGTGAGCGCCGACGACAATGTCTCATTAACGGTGGGGTAAGCCTTCG        720

Qy    721  GAGATATTTCTCCTGCAACAAGCACAGGGAAATGCCGGACGGGAAGAGC          780
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  GAGATATTTCTCCTGCAACAAGCACAGGGAAATGCCGGACGGGAAGAGC         780

Qy    781  GATTCACACCAGTGCGAACACCTTGCTAAGTTTGCATAACGGGCAAATTT        840
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  GATTCACACCAGTGCGAACACCTTGCTAAGTTTGCATAACGGGCAAATTT       840

Qy    841  CGCACGCCAGAGGTTGCCCGCGAGCGCGCCACCCCGGTTATTGGATTTG         900
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  CGCACGCCAGAGGTTGCCCGCGAGCGCGCCACCCCGGTTATTGGATTTG        900

Qy    901  TTGACGCCCCATCCACCGCAAAAAACAGGGCGTATGGTGTGACATTACC         960
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  TTGACGCCCCATCCACCGCAAAAAACAGGGCGTATGGTGTGACATTACC        960

Qy    961  TTTATTGCGGACACGAATACTAATCTCGGCGGCAAATCTCGGCGGCAACTCGG     1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  TTTATTGCGGACACGAATACTAATCTCGGCGGCAAATCTCGGCGGCAACTCGG    1020

Qy   1021  ACGGCTTCCCGGTCAGCCGGGATAACACGCGCCAGGTGGTGAACTGGGTTG       1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  ACGGCTTCCCGGTCAGCCGGGATAACACGCGCCAGGTGGTGAACTGGGTTG      1080

Qy   1081  CGTCGGGCTAAGCGATAACAGCCAGTGAGATTCAGGTTTCGGCTGCGTTACAG      1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  CGTCGGGCTAAGCGATAACAGCCAGTGAGATTCAGGTTTCGGCTGCGTTACAG     1140

Qy   1141  CAGATGCCTGATAAAACGCCGCTGTCATTAAAATACGCGCCCGAGAGGTGAAACTGACC     1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  CAGATGCCTGATAAAACGCCGCTGTCATTAAAATACGCGCCCGAGAGGTGAAACTGACC    1200

Qy   1201  CTGGCAGGATGTGAAGAGGCGAAATGCCGGCCATGTGTTCGGTTGGCAGGTTTACGCAA    1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  CTGGCAGGATGTGAAGAGGCGAAATGCCGGCCATGTGTTCGGTTGGCAGGTTTACGCAA   1260

Qy   1261  ATCGTGAATGAAGCACGCATACCCGGCGGGTGCGTGCAGTTTGAGATTCATCACCATCAC   1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261  ATCGTGAATGAAGCACGCATACCCGGCGGGTGCGTGCAGTTTGAGATTCATCACCATCAC  1320

Qy   1321  TAA  1323
           |||
Db   1321  TAA  1323
```

RESULT 6
US-09-866-379-7

; Sequence 7, Application US/09866379
; Patent No. 6855365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 7
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-7

Query Match   97.4%; Score 1287.8; DB 3; Length 1901;
Best Local Similarity   99.5%; Pred. No. 0; Mismatches 7; Indels 0; Gaps 0;
Matches 1293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy   1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAAACCCCGAAATCTGCA 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAAACCCCGAAATCTGCA 247

Qy   61 TTCGCTCAGAGTGAGCCGAGCTGAAAGTGGTGATTGTCAGTCGTCATTGT... 120
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db   248 TTCGCTCAGAGTGAGCCGAGCTGAAAGTGGTGATTGTCAGTCGTCATGGT 307

Qy   121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGATGCAGAGATGTCACCCCGAGACGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGATGCAGAGATGTCACCCCGAGACGCCA 367

Qy   181 ACCTGGCCGCGGGTAAAACTGGGTTGGGCTGCACACCGCCGGCGGATTGCTGGCGAAAAG 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368 ACCTGGCCGCGGGTAAAACTGGGTTGGGCTGCACACCGCCGGCGGATTGCTGGCGAAAAG 427

Qy   241 GGACATTACCAACGCCAGCGTCTGGGTAGCCGACGGATTGCTGGCGGAAAAAGGGCTGCCCG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 GGACATTACCAACGCCAGCGTCTGGGTAGCCGACGGATTGCTGGCGGAAAAAGGGCTGCCCG 487

Qy   301 CAGTCTGGTCAGGTCAGGTCGGCGATTATTGCTGAATGTCGACGAGGCGGA 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488 CAGTCTGGTCAGGTCAGGTCGGCGATTATTGCTGAATGTCGACGAGGCGGA 547

Qy   361 GCCTTCGCCGCCGGGCCGGGGCTGGGCACCTGACTGTGCAATACCGAGGATACC 420
              |||||||||||||||||||||||||||||||||||||||||||||||
Db   548 GCCTTCGCCGCCGGGCCGGGGCTGGGCACCTGACTGTGCAATACCGAGGATACG 607

Qy   421 TCCAGTCCCGATCCGGTTATTTATTTAATCCTTCTAAAAAACTGGCGTTTTACGGCCG 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   608 TCCAGTCCCGATCCGGTTATTTATTTAATCCTTCTAAAAAACTGGCGTTTTACGGCCAT 667

Qy   481 AACGTGACTGACCGGATCCTCAGCACGGGGCCAGGAGGGGTCAATGCTGACG 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668 AACGTGACTGACCGGATCCTCAGCACGGGGCCAGGAGGGGTCAATGCTGACG 727

Qy   541 CGGCAAAACGGGCGTTTCCGGAACTGGAACGGGCAGGGGGTGGGGAGGGGAAGGGAACTC 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   728 CGGCAAAACGGGCGTTTCCGGAACTGGAACGGGCAGGGGGTGGGGAGGGGAAGGGAACTC 787

Qy   601 CTTAAACGTGAGAAAACAGGACGAAAGCTTGTTCATTAAACGCAGGCAATTACCATCGGGAACTC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   788 CTTAAACGTGAGAAAACAGGACGAAAGCTTGTTCATTAAACGCAGGCAATTACCATCGGGAACTC 847

Qy   661 AAGGTGAGCCGCCGACAAATGTCTCATTAACCGGTGGCGGTAAGCCTCGGATCAATGCTGACG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   848 AAGGTGAGCCGCCGACAAATGTCTCATTAACCGGTGGCGGTAAGCCTCGGATCAATGCTGACG 907

Qy   721 GAGATATTTCTCCTGCAACAAGCACACGGAAGTGCTTAATTTTCCGCAATCAAAACTTGTGC 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   908 GAGATATTTCTCCTGCAACAAGCACACGGAAGTGCTTAATTTTCCGCAATCAAAACTTGTGC 967

Qy   781 GATTCACCAGTGGAACACCTTGCTTAAGTTTGCATAAACGGCAATTTATTTGCTACAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db   968 GATTCACCAGTGGAACACCTTGCTTAAGTTTGCATAAACGGCAATTTATTTGCTACAA 1027

Qy   841 CGCACGCCAGAGTTGCCCGCAGCCGCGGGCCACCCCGTTATTGGATTGGCAGCG 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028 CGCACGCCAGAGTTGCCCGCAGCCGCGGGCCACCCCGTTATTGGATTGGCAGCG 1087

Qy   901 TTGACGCGCCCATCCACCGCAAAAAACAGGCGTATGGGTGGTGGACATTACCGACT 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 TTGACGCGCCCATCCACCGCAAAAAACAGGCGTATGGGTGGTGGACATTACCGACTGCTG 1147

Qy   961 TTTATTGCCGGACACGATACTAATCTGGCAAAATCTCGGCGGCCACTGGAGCTCAACTGG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148 TTTATTGCCGGACACGATACTAATCTGGCAAAATCTCGGCGGCCACTGGAGCTCAACTGG 1207

Qy   1021 ACGGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGGTCGAACCGGTGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208 ACGGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGGTCGAACCGGTGG 1267

Qy   1081 CGTCGGCGTAAGCCGCTACAGCAGTGGATTCAGGTTTCGCTGCTGGTCTTCCAGACTTTACAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268 CGTCGGCGTAAGCCGCTACAGCAGTGGATTCAGGTTTCGCTGCTGGTCTTCCAGACTTTACAG 1327

Qy   1141 CAGATGCGTGATTAAAAACGCCCGTTCATTAAAATACCGCGCCGGTGAAAACTGACC 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328 CAGATGCGTGATTAAAAACGCCCGTTCATTAAAATACCGCGCCGGTGAAAACTGACC 1387

Qy   1201 CTGGCAGGGATGTGAAAGAGGCGAAAATGCGCAGGGCCATGTGTTCGGTTGGCAGGGTTTTACGGCAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388 CTGGCAGGGATGTGAAAGAGGCGAAAATGCGCAGGGCCATGTGTTCGGTTGGCAGGGTTTTACGGCAA 1447

Qy   1261 ATCGTGAATGAATGAAGCACGCATACCGGCGCTGCAGTTTGACAT 1300
              ||||||||||||||||||||||||||||||||||||||||||
Db   1448 ATCGTGAATGAATGAAGCACGCCATACCGGCGCTGCAGTTTGTAAT 1487

RESULT 7
US-09-866-379-5
; Sequence 5, Application US/09866379

Qy   1201 CTGGCAGGATGTGAAGAGCAAGCGCATACCGGGTCATGTTTCGGTGGCAGGGTTTTACGGCAA 1260
              |||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAGCACGCAATACCGGGTCATCATCACCATCACCATCAC 1320

Qy   1261 ATCGTGAATGAAGCACGCAATACCGGGTGTCATCACCATCACCATCAC 1320
              ||||||||||||||||||||||||||||||||||||||||||
Db   1321 TAA 1323

Qy   1321 TAA 1323
              |||
Db   1321 TAA 1323

```
; Patent No. 6855365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
;
; SEQ ID NO 5
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-5

  Query Match             97.3%;  Score 1286.2;  DB 3;  Length 1901;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1292;  Conservative 0;  Mismatches 8;  Indels 0;  Gaps 0;

Qy         1 ATGAAAACGCATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       188 ATGAAAACGGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 247

Qy        61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAAGTGTGGTAGTCAGTCGTCGTCATGGT 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       248 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAAGTGTGGTAGTCAGTCGTCGTCATGGT 307

Qy       121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATCCAGAGATGTCACCCCGACACGGTCCAGG 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATCCAGAGATGTCACCCCGACACGGTCCAGG 367

Qy       181 ACCTGGCCGGGTAAAACTGGGGTTGGCCTGACATAATCGGCCTATCTC 240
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db       368 ACCTGGCCGGGTAAAACTGGGGTTGACTGACACCCCGCGNNGTGGCTAATCGGCCTATCTC 427

Qy       241 GGACATTACCAACGCCAGCGTCTGGGTAGCCGACGAGGATTGCTCGGCAGAAAAAGGGCTGCCCG 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       428 GGACATTACCAACGCCAGCGTCTGGGTAGCCGACGAGGATTGCTCGGCAGAAAAAGGGCTGCCCG 487

Qy       301 CAGTCTCGTCAGGTCGGCCATTATTGCTCTGATGTCGACAGGCGTACCCGTAAAAACAGGCGAA 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       488 CAGTCTCGTCAGGTCGGCGAATTATTGCTCTGATGTCGACAGGCGTACCCGTAAAAACAGGCGAA 547

Qy       361 GCCTTCGCCGCCGCGGGCTGGCAACTGACCGTGCAATAACCGTACATACCCGGCAGATACG 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       548 GCCTTCGCCGCCGCGGGCTGGCAACTGACCGTGCAATAACCGTACATACCCGGCAGATACG 607

Qy       421 TCCAGTCCCGATCCGGTTATTTAATCCTCTAAAAAACTGGCGGTTGCCAACTGGATAACGGCG 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       608 TCCAGTCCCGATCCGGTTATTTAATCCTCTAAAAAACTGGCGGTTGCCAACTGGATAACGGCG 667

Qy       481 AACGTGACTGACCGGATCCTCAGCAGGGCAGAGGGTCAATTGCTGACTTTACCGGGGCAT 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       668 AACGTGACTGACCGGATCCTCAGCAGGGCAGAGGGTCAATTGCTGACTTTACCGGGGCAT 727

Qy       541 CGGCAAAACGGCGTTTCGGCGAACTGGAAACGGCGGTGCTTAATTTTCCCGCAATCAAACTTGTGC 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       728 CGGCAAAACGGCGTTTCGGCGAACTGGAAACGGCGGTGCTTAATTTTCCCGCAATCAAACTTGTGC 787

Qy       601 CTTAAACGTTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       788 CTTAAACGTTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC 847

Qy       661 AAGGTGAGCGCCCGACAAATGTCTCATTAACCCGGTGCGGTAAGCCTCGCCTCGCCATCAAATGCTGACG 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       848 AAGGTGAGCGCCCGACAAATGTCTCATTAACCCGGTGCGGTAAGCCTCGCCTCGCCATCAAATGCTGACG 907

Qy       721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGTTGGGGAAGGGATCACC 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       908 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGTTGGGGAAGGGATCACC 967

Qy       781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCGCAATTTATTTATTTGCTACAA 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       968 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCGCAATTTATTTATTTGCTACAA 1027

Qy       841 CGCACGCCAGAGGTTGCCCCGCCAGCCCGCCACCCCGTTATTAGGATTTGATCATGGCACGCG 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1028 CGCACGCCAGAGGTTGCCCCGCCAGCCCGCCACCCCGTTATTAGGATTTGATCATGGCACGCG 1087

Qy       901 TTGACGCCCCATCCACCCGCAAAAACAGGCGTATGGTGTGACATTACCACTTCAGTGCTG 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1088 TTGACGCCCCATCCACCCGCAAAAACAGGCGTATGGTGTGACATTACCACTTCAGTGCTG 1147

Qy       961 TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCCGCACTGGAGGCTCAACTGG 1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1148 TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCCGCACTGGAGGCTCAACTGG 1207

Qy      1021 ACGCTTCCCGGTCAGCCCGGATAACACGCCCGCCGGTGGGTGGTGTTGAAGCCTAC 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1208 ACGCTTCCCGGTCAGCCCGGATAACACGCCCGCCGGTGGGTGGTGTTGAAGCCTAC 1267

Qy      1081 CGTCGGCTAAGCGGATAACAGCCACTGGGATTCAGGTTTCGTCTTCGCTGGGTCTTCAGACTTTACAG 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1268 CGTCGGCTAAGCGGATAACAGCCACTGGGATTCAGGTTTCGTCTTCGCTGGGTCTTCAGACTTTACAG 1327

Qy      1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATAACGGCGGGTTCGGGGAGAGGTGAAACTGACC 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1328 CAGATGCGTGATAAAACGCCGCTGTCATTAAATAACGGCGGGTTCGGGGAGAGGTGAAACTGACC 1387

Qy      1201 CTGGCAGGATTGTGAAGAGCCGAAATGCCAGGGCAGTGTTCGTTGGGCAGTTTACGGCAA 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1388 CTGGCAGGATTGTGAAGAGCCGAAATGCCAGGGCAGTGTTCGTTGGGCAGTTTACGGCAA 1447

Qy      1261 ATCCTGAATGAAGCACGCCATACCCGGCGTGCAGTTGTGTAAT 1300
             ||||||||||||||||||||||||||||||||||||||||||
Db      1448 ATCCTGAATGAAGCACGCCATACCCGGCGTGCAGTTGTGTAAT 1487
```

```
RESULT 8
US-09-866-379-9
; Sequence 9, Application US/0986379
; Patent No. 6855365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
```

PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentin version 3.1

SEQ ID NO 9
  LENGTH: 1901
  TYPE: DNA
  ORGANISM: Escherichia coli
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (1)..(1901)
    OTHER INFORMATION: n is any nucleotide

US-09-866-379-9

Query Match   97.3%; Score 1286.2; DB 3; Length 1901;
Best Local Similarity   99.4%; Pred. No. 0; Mismatches 8; Indels 0; Gaps 0;
Matches 1292; Conservative 0;

```
Qy    1 ATGAAAAGCGGATCTTAATCCCATTTTATCTCTTCTCTGAATTCCGTTAACCCCGCAATCTGCA   60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  188 ATGAAAGCGGATCTTAATCCCATTTTATCTCTTCTCTGAATTCCGTTAACCCCGCAATCTGCA  247

Qy   61 TTCGCTCAGAGTGAGCCGGGAGCTGAAGCTGGAAAGTGTGGAATTGTCAGTCGTCGTTCATGGT  120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  248 TTCGCTCAGAGTGAGCCGGGAGCTGAAGCTGGAAAGTGTGGAATTGTCAGTCGTCGTTCATGGT  307

Qy  121 GTGCGTGCTCCAACCAAGGCACGACGCAACTGATGCAGGAGCAGCCGCCATGGCCA  180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  308 GTGCGTGCTCCAACCAAGGCCACGACGCAACTGATGCAGGAGCAGCCGCCATGGCCA  367

Qy  181 ACCTGGCGCGGTAAAACTGGGTTGGGCTGCAGCGGCTAAATCGCCTATCTC  240
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  368 ACCTGGCGCGGTAAAACTGGGTTGGGCTGCAGCGGCTAAATCGCCTATCTC  427

Qy  241 GGACAATTACCAAACGCCAGCGTCTGGGTAGCCGACGGAATTGCTTGGCGAAAAAGGGCTGCCCG  300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  428 GGACAATTACCAAACGCCAGCGTCTGGGTAGCCGACGGAATTGCTTGGCGAAAAAGGGCTGCCCG  487

Qy  301 CAGTCTCGGTCAGGTCGGGCGATTATTGCTGATGTCGACACGGATACCCGTAAAACAGGCGAA  360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  488 CAGTCTCGGTCAGGTCGGGCGATTATTGCTGATGTCGACACGGATACCCGTAAAACAGGCGAA  547

Qy  361 GCCTTGCCGCCGCCGGCGTGGCACCTGACTGTGCAATAACCGTACATACCCAGCCAGATACG  420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  548 GCCTTCGCCGCCGGCGTGGCACCTGACTGTGCAATAACCGTACATACCCAGCCAGATACG  607

Qy  421 TCCAGTCCCGATCCGGTTATTTAATCCTTCTAAAAAACTGGATAACGCG  480
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  608 TCCAGTCCCGATCCGGTTATTTAATCCTTCTAAAAAACTGGATAACGCG  667

Qy  481 AACGTGACTGAGCGGATCCTCAGCAGGCAGAGGGGTTCAAATTGCTGACTTTACCGGGGCAT  540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  668 AACGTGACTGAGCGGATCCTCAGCAGGCAGGAGGGGTTCAAATTGCTGACTTTACCGGGGCAT  727

Qy  541 CGGCAAACGGGTTTGCGGACACTGGAACGGGGTGCTTAATTTCCGCAAATCAAACTTGTGC  600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  728 CGGCAAACGGGTTTGCGGACACTGGAACGGGGTGCTTAATTTTCCGCAAATCAAACTTGTGC  787

Qy  601 CTTAAACGTGAGAAAACAGGACGAAACGTGTTCATTAACGCAGGCAATTACCATCGGAACTC  660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  788 CTTAAACGTGAGAAAACAGGACGAAACGTGTTCATTAACGCAGGCAATTACCATCGGAACTC  847

Qy  661 AAGGTGAGCGCCGACAATGTCTCATTAACCGGGTGCGGTAAGCCTCGGCATCAATGCTGACG  720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  848 AAGGTGAGCGCCGACAATGTCTCATTAACCGGGTGCGGTAAGCCTCGGCATCAATGCTGACG  907

Qy  721 GAGATATTTCCTGCAACAGCACAGGGAATGCCGGAGCCGGGGTGGGGAGGATCACC  780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      (780)
```

RESULT 9

US-09-866-379-6

Sequence 6, Application US/09866379
Patent No. 6855365

GENERAL INFORMATION:
  APPLICANT: DIVERSA CORPORATION
  APPLICANT: SHORT, Jay
  APPLICANT: KRETZ, Keith
  APPLICANT: GRAY, Kevin
  APPLICANT: BARTON, Nelson
  APPLICANT: GARRETT, James
  APPLICANT: O'DONOGHUE, Eileen
  TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
  FILE REFERENCE: DIVER1370-7
  CURRENT APPLICATION NUMBER: US/09/866,379
  CURRENT FILING DATE: 2001-05-24
  PRIOR APPLICATION NUMBER: US 09/580,515
  PRIOR FILING DATE: 2000-05-25
  PRIOR APPLICATION NUMBER: US 09/318,528
  PRIOR FILING DATE: 1999-05-25
  PRIOR APPLICATION NUMBER: US 09/291,931
  PRIOR FILING DATE: 1999-04-13
  PRIOR APPLICATION NUMBER: US 09/259,214
  PRIOR FILING DATE: 1999-03-01
  PRIOR APPLICATION NUMBER: US 08/910,798
  PRIOR FILING DATE: 1997-08-13
  NUMBER OF SEQ ID NOS: 10
  SOFTWARE: Patentin version 3.1

SEQ ID NO 6
  LENGTH: 1901
  TYPE: DNA
  ORGANISM: Escherichia coli
  FEATURE:

```
Db  908 GAGATATTTCTCCTGCAACAGCACAGGGAATGCCGGAGCCGGGGTGGGGAGGATCACC  967

Db  968 GATTCACACCAGTGGGAACACCTTGCTAAGTTTGCATAACGCCCAATTTTATTTGCTACAA  1027

Db 1028 CGCACGCCAGAGGTTGCCCGCGCCACCCCGTTATTGGATTGAATCAAGACAGCG  1087

Db 1088 TTGACGCCCCATCCACCCGCAAAAAACAGGCGTATGGGTGTTGACATTACCCACTTCAGGTGCTG  1147

Db 1148 TTTATCGCCGGACACAGCGATACTAATCCTGGCAAATCTCGGCGGCTCAACTGG  1207

Db 1208 ACGCTTCCCGGTCAGCCCGGATAAACACCGCCCAGGTGGTGAACTGGGTGTTGAACGGCTGG  1267

Db 1268 CGTCGGCTAAGCCGATAACAGCCAGTGGGAATTCAGGTTCGCGTCGGTTCTTCCAGACTTTACAG  1327

Db 1328 CAGATGCGTGATAAAACGCCGTGTCATTAAAATACGGGCCCCGGGAGAGGTGAAACTGACC  1387

Db 1388 CTGGGCAGGATGTGAAGAGACGGAAAATGCCGAGGGCATGTGGTTTCCGTTGGCGAGGTTTTTACCGCAA  1447

Db 1448 ATCGTGAATGAAGCACACGGCGATACCGGGCGTGCCAGTTTGTAAT  1487
```

```
; NAME/KEY: misc_feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-6

Query Match    97.0%;  Score 1283;  DB 3;  Length 1901;
Best Local Similarity   99.2%;  Pred. No. 0;
Matches 1290;  Conservative  0;  Mismatches  10;  Indels  0;  Gaps  0;

Qy      1 ATGAAAAGCGATCTTAATCCATTTTATTCTTCTGATTCCGTTAACCCCGCAAATCTGCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188 ATGAAAAGCGATCTTAATCCCATTTTATCTCTTCCGTTAACCCCGCAAATCTGCA 247

Qy     61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTTGAAAGTGTTGTCAGTCGTCATGGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTTGAAAGTGTTGTCAGTCGTCATGGT 307

Qy    121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 367

Qy    181 ACCTGGCCGGTAAAACTGGGTTGGCTGACAACCGGTAATCGCGCCTATCTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    368 ACCTGGCCGGTAAAACTGGGTTGGCTGACAACCGGTAATCGCGCCTATCTC 427

Qy    241 GGACATTACCAACGCCAGCGTTCTGGTAGCCGACAGGATTGCTGGGGAAAAGGGCTGCCCG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428 GGACATTACCTGGGCCACGCGTTCTGGTAGCCGACAGGATTGCTGGGGAAAAGGGCTGCCCG 487

Qy    301 CAGTCTCGGTCAGGTCGCGCGATTATTGCTGATGTCGACGAGGTACCCGTAAAACAGGCGAA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488 CAGTCTCGGTCAGGTCGCGGATTATTGCTGATGTCGACGAGGTACCCGTAAAACAGGCGAA 547

Qy    361 GCCTTCGCCCGCGCCGGGCTGGACATGTGACACTGACTACCAGGCAGGATACG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548 GCCTTCGCCCGCGCCGGGCTGGACATGTGACACTGACTACCAGGCAGGATACG 607

Qy    421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAAACTGGCGTTTGCCAACTGGATAACGCG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAAACTGGCGTTTGCCAACTGGATAACGCG 667

Qy    481 AACGTGACTGACCGGATCCTCAGCGAGGGCCAGGAGGGTCAAATTGCTGACTTTACCGGCAT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    668 AACGTGACTGACCGGATCCTCAGCGAGGGCCAGGAGGGTCAAATTGCTGACTTTACCGGGCAT 727

Qy    541 CGGCAAAACGGCGTTTCCGGAACTGGAACGGGGTGCTTAAATTTCCGGCAATCAAACTTGTGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    728 CGGCAAAACGGCGTTTCCGGAACTGGAACGGGGTGCTTAATTTTCCGGCAATCAAACTTGTGC 787

Qy    601 CTTAAACGTGAGAAACAGGAAGCGAAAGCTGTTCATTAACGGCAGGCAATTACCCATCCATGGAACTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    788 CTTAAACGTGAGAAACAGGAAGCGAAAGCTGTTCATTAACGGCAGGCAATTACCCATCCATGGAACTC 847

Qy    661 AAGGTGAGCGCCCGACAATGTCTCATTAACCGGTGCGGTAAGCTCGCATCAAATGCTGACG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    848 AAGGTGAGCGCCCGACAATGTCTCATTAACCGGTGCGGTAAGCTCGCATCAAATGCTGACG 907

Qy    721 GAGATATTTCTCCTGCAACAAGCACAGGGAATCCGGACGGATGATGTCACCCAGACGGCCA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908 GAGATATTTCTCCTGCAACAAGCACAGGGAATCCGGACGGATGATGTCACCCAGACGGCCA 967

Qy    781 GATTCAACACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCGCAATTTATTGCTACAA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968 GATTCAACACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCGCAATTTATTGCTACAA 1027

Qy    841 CGCACGCCAGAGGTTGCCCGCGCCCCGGTTATTGGATTTGGATTTCATGGCAGCG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028 CGCACGCCAGAGGTTGCCCGCGCCCCGGTTATTGGATTTGGATTTCATGGCAGCG 1087

Qy    901 TTGACGCCCCATCCACGCGGCCAAAAAACAGGCCGTAGGCGTGTGACATTACCCACTTCAGTGCTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 TTGACGCCCCATCCACGCGGCCAAAAAACAGGCCGTAGGCGTGTGACATTACCCACTTCAGTGCTG 1147

Qy    961 TTTATTGCCGGACACGATTACTAATCCTGGCAAAATCTCGGCGGCCGACACTGGGAGCTCAACTGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148 TTTATCGCCGGACACGATTACTAATCCTGGCAAAATCTCGGCGGCCGACACTGGGAGCTCAACTGG 1207

Qy   1021 ACGCTTCCCGTCAGCCGGATAAACAGCCGGCTAGACTAAACGCCCGGCAGGTTGAAACGGTGG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208 ACGCTTCCCGTCAGCCGGATAAACAGCCGGCTAGACTAAACGCCCGGCAGGTTGAAACGGTGG 1267

Qy   1081 CGTCCGCTAAGCGATAACAGCCAGTCGGATTCAGGTTCGCGGTCTTCCAGACTTTACAG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268 CGTCCGCTAAGCGATAACAGCCAGTCGGATTCAGGTTCGCGGTCTTCCAGACTTTACAG 1327

Qy   1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATAACGGCCCGCGCGGAGGGTGAAACTGACC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328 CAGATGCGTGATAAAACGCCGCTGTCATTAAATAACGGCCCGCGCGGAGGGTGAAACTGACC 1387

Qy   1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCCATGTGTTCGTTGGCAGGTTTTACGGCAA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCCATGTGTTCGTTGGCAGGTTTTACGGCAA 1447

Qy   1261 ATCGTGAATGAAGCACGCAATACCGCGCGTGCAGTTTGAGAT 1300
          ||||||||||||||||||||||||||||||||||||||||||
Db   1448 ATCGTGAAATGAAGCACGCAATACCGCGCGTGCAGTTTGTAAT 1487
```

```
RESULT 10
US-09-540-149A-9
; Sequence 9, Application US/0954014A9
; Patent No. 651699
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2791
; CURRENT APPLICATION NUMBER: US/09/540,149A
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 60/127,032
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1489
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-540-149A-9

Query Match    95.6%;  Score 1263.8;  DB 3;  Length 1489;
Best Local Similarity   97.9%;  Pred. No. 0;
Matches 1280;  Conservative  0;  Mismatches  28;  Indels  0;  Gaps  0;

Qy      1 ATGAAAAGCGATCTTAATCCATTTTATTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    182 ATGAAAAGCGATCTTAATCCCATTTTATCTCTTCCGTTAACCCCGCAAATCTGCA 241

Qy     61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTTGAAAGTGTTGTCAGTCGTCATGGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    242 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTTGAAAGTGTTGTCAGTCGTCATGGT 301

Qy    121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    302 GTGCGTGCCCCAACCAAGGCCACCGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 361

Qy    181 ACCTGGCCGGTAAAACTGGGTTGGCTGACAACCGGTAATCGCGCCTATCTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    362 ACCTGGCCGGTAAAACTGGGTTGGCTGACAACCGGTAATCGCGCCTATCTC 421

Qy    241 GGACATTACCAACGCCAGCGTTCTGGTAGCCGACAGGATTGCTGGGGAAAAGGGCTGCCCG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    422 GGACATTACCAACGCCAGCGTTCTGGTAGCCGACAGGATTGCTGGGGAAAAGGGCTGCCCG 481

Qy    301 CAGTCTCGGTCAGGTCGCGGATTATTGCTGATGTCGACGAGGCGTACCCGTAAAACCGGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    482 CAGTCTCGGTCAGGTCGCGGATTATTGCTGATGTCGACGAGGCGTACCCGTAAAACCGGAA 541
```

RESULT 11
US-10-266-041A-9
Sequence 9, Application US/10266041A
Patent No. 6974690
GENERAL INFORMATION:
   APPLICANT: Lei, Xingen
   TITLE OF INVENTION: PHOSPHATASESS WITH IMPROVED PHYTASE ACTIVITY
   FILE REFERENCE: 19603/2798

; CURRENT APPLICATION NUMBER: US/10/266,041A
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: 60/127,032
; PRIOR FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 09/540,149
; PRIOR FILING DATE: 2000-03-31
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 9
;   LENGTH: 1489
;   TYPE: DNA
;   ORGANISM: Escherichia coli
; US-10-266-041A-9

Query Match   95.6%;   Score 1263.8;   DB 4;   Length 1489;
Best Local Similarity 97.9%;   Pred. No. 0;   Mismatches 28;   Indels 0;   Gaps 0;
Matches 1280; Conservative   0;   Mismatches   28;   Indels   0;   Gaps   0;

[dense nucleotide sequence alignment blocks labeled Qy (query) and Db (database), with positional numbering, spanning the full page]

RESULT 12

US-09-715-477-2

; Sequence 2, Application US/09715477
; Patent No. 6841370
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
; FILE REFERENCE: 19603/4031
; CURRENT APPLICATION NUMBER: US/09/715,477
; CURRENT FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1486
; TYPE: DNA
; ORGANISM: Escherichia coli

US-09-715-477-2

Query Match   95.5%;   Score 1263;   DB 3;   Length 1486;
Best Local Similarity   98.4%;
Matches 1275; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy   1 ATGAAAGCGATCTTAATCCCATTTTATCTCTTCGATTCCGTTAACCCGCAAATCTGCA   60

Db   188 ATGAAAGCGATCTTAATCCCATTTTATCTCTTCGATTCCGTTAACCCGCAAATCTGCA   247

Qy   61 TTCGCTCAGAGTGAGCCCGAGCTGAAGCTGGAAAAGTGTGGTGATTGTCAGTCGTCATGGT 120

Db   248 TTCGCTCAGAGTGAGCCCGAGCTGAAGCTGGAAAAGTGTGGTGATTGTCAGTCGTCATGGT 307

Qy   121 GTGGCGCTCTCAACCAACAAGGGCCCACGCGCCACTGACTGATCGCAGATGGCGGCATGGCCA 180

Db   308 GTGGCGCTGCCCAACCACGGCCACGCTGATCGACGAGTGCACCCCAGACCGCATGGCCA 367

Qy   181 ACCTGGCGCGGTAAAACTGGGTTGGCTGACACCGGCGGTGAGCGCTAAATCGGCCTATCTC 240

Db   368 ACCTGGCGCGGTAAAACTGGGTTGGCTGACACCGGCGGTGAGCGCTAAATCGGCCTATCTC 427

Qy   241 GGACATTACCAAACGCCAGGTGTCTGGTTGCTGGCCGGACGCGGAT TGCTGGCCGGAC   300

Db   428 GGACATTACCAAACGCCAGGTGTCTGGTTGCTGGCCGGACGCGGATTGCTGGCCGGAC   487

Qy   301 CAGTCTTGGTTCAGGTTCGGCGAT TCAGCGGACGGAGCGGTACCCGTAAAACAGG   360

Db   488 CAGCCTTGGTTCAGGTTCGGCGAT TCAGCGGACGGAGCGGTACCCGTAAAACAGG   547

Qy   361 GCCTTCGCCGCCGGCCTGGGCTGGCCACCTGACCTGTGCAATAACCGTACATACCCAGG   420

Db   548 GCCTTCGCCGCCGGCCTGGGCTGGCCACCTGACCTGTGCAATAACCGTACATACCCAGG   607

Qy   421 ACGGCTTCCCGGTCAGCCGGATAAACACGCCGCCCAGGTGGTGAACTGGCTGTTTGAACGGCTGG 480

Db   608 TCCAGTCCCGGATCCCGTTAATTTAATCCCTCTAAAAAACTGGCCGGTTGCCAACTGGATAACGCGG 667

Qy   481 AACCTTGACTTGACCGGGATCCTTCAGCAGGGGGTCAAATTGCTGACTTTTACCCGGGCAT 540

Db   668 AACCTTGACTTGACCGGGATCCTTCAGCAGGGGGTCAAATTGCTGACTTTTACCCGGGCAT 727

Qy   541 CGGCAAAACGGGCGTTTTCGCGAACTGGAACGGGGTGCTTAATTTCCGGCAACTGGTTGC 600

Db   728 CGGCAAAACGGGCGTTTTCGCGAACTGGAACGGGGTGCTTAATTTCCGGCAACTGGTTGC 787

Qy   601 CTTAAACGTGAGAAAACAGGACGAAAAGCTGTTCATTAACCGGGAACTCGAACTC 660

Db   788 CTTAAACGTGAGAAAACAGGACGAAAAGCTGTTCATTAACCGGGAACTCGAACTC 847

Qy   661 AAGGTGAGCGCCGACAAATGTCTCATTAACCGGTCGCGGTAAGCCTCGCATCAAATGCTGACG 720

Db   848 AAGGTGAGCGCCGACAAATGTTTCATTAACCGGTCGCGGTAAGCCTCGCATCAAATGCTGACG 907

Qy   721 GAGATATTTCTCCTGCAACAAGCACAGGGAAGTGCCGGACCGGGGGGTGGGGGAAGGATCACC 780

Db   908 GAAATATTTCTCTGCAACAAGCACAGGGAAGTGCCGGACCGGGGGGTGGGGGAAGGATCACT 967

Qy   781 GATTCACACCAGTGGGAAACACCTTGCTTGCATAGTTTGCATAACGGCAATTTATTTGCTACAA 840

Db   968 GATTCACACCAGTGGGAAACACCTTGCTTGCATAGTTTGCATAACGGCAATTTATTTACTACAA 1027

Qy   841 CGCACGCCAGAGGTTGCCCGGCACGCGCGCCACCCCGTTATTGGATTTGATCATGGCAGCG 900

Db   1028 CGCACGCCAGAGGTTGCCCGGCACGCGCGCCACCCCGTTATTGGATTTGATCAACAGAGCG 1087

Qy   901 TTGACGCCCCATCCACCGGCAAAAACAGGGCGTATGGGTGGACATTACCCCACTTCAGTGCTG 960

Db   1088 TTGACGCCCCATCCACCGGCAAAAACAGGGCGTATGGGTGGACATTACCCCACTTCAGTGCTG 1147

Qy   961 TTTATTGCCGGACACGATACTAATCTGGCAAAATCTCGGGCGGCAGCTCAACTGG 1020

Db   1148 TTTATTGCCGGACACGATACTAATCTGGCAAAATCTCGGGCGGCAGCTCAACTGG 1207

Qy   1021 ACGGCTTCCGGGTCAGCCGGATAACACGCCGCCCAGGTGGTGAACTGGCTGTTTGAACGGCTGG 1080

Db   1208 ACGGCTTCCAGGTCAGCCGGATAACACGCCGCCCAGGTGGTGAACTGGCTGTTTGAACGGCTGG 1267

Qy   1081 CGTCGGCTAAGCGGATAAACAGCCAGTGGAATTCAGGTTTCGGCTCTTCCAGACTTTACAG 1140

Db   1268 CGTCGGCTAAGCGGATAAACAGCCAGTGGAATTCAGGTTTCGGCTCTTCCAGACTTTACAG 1327

Qy   1141 CAGATGCGTGATAAAACCCGCTGTCATTAAAATACGGCCGCCCGGGAAAACTGACC 1200

Db   1328 CAGATGCGTGATAAAACCCGCTATCATTAAAATACGGCCGCCCGGGAAAACTGACC 1387

Qy   1201 CTGGCAGGATGTGAAGAGGCAGGGCATGTGTTGGGCGGTTTTACGGCA 1260

Db   1388 CTGGCAGGATGTGAAGAGGCAGGGCATGTGTTGGGCGGTTTTACGGCAA 1447

Qy   1261 ATCGTGAATGAAGGACGCGCATACCCGGCGTTGCAGTTTG 1296

Db   1448 ATCGTGAATGAAGGACGCGCATACCCGGCGTTGCAGTTTG 1483

RESULT 13
US-09-715-477-4
; Sequence 4, Application US/09715477
; Patent No. 6841370
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
; FILE REFERENCE: 19603/4031
; CURRENT APPLICATION NUMBER: US/09/715,477
; CURRENT FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
;
; SEQ ID NO 4
; LENGTH: 1486
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-715-477-4

Query Match 94.7%; Score 1251.8; DB 3; Length 1486;
Best Local Similarity 97.8%; Pred. No. 0; Mismatches 28; Indels 0; Gaps 0;
Matches 1268; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy   1 ATGAAAAGCGATCTTAATCCATTTTAATCCTTAACCCCGGAAATCTGCA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 188 ATGAAAAGCGATCTTAATCCCATTTTAATCCTTAACCCCGGAAATCTGCA 247

Qy   61 TTCGCTCAGAGTGAGCCCGGAGCTGAAGCTGGAGTGTGGTGATTGTCAGTTGTCATGGT 120
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 248 TTCGCTCAGAGTGAGCCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTTGTCATGGT 307

Qy 121 GTGGCGGTGCTCAACCAAGGCCACGCAACTGATAGCGATGTCACCCCAGAGGCATGGCCA 180
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 308 GTGGCGTGCCCAACCAAGGCCACGCAACTGATGCAGATGTCACCCCAGACGCATGGCCA 367

Qy 181 ACCTGGCCGGTAAAAACTGGGTGGTTGGCTGACACCGCGGGTGAGCTAATCGCCTAITCTC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 368 ACCTGGCCGGTAAAACTGGGTGGTTGGCTGACACACCGCGGGTGAGCTAATCGGCCTAITCTC 427

Qy 241 GGACAITACCAACGCCAGCGTCTGGTAGCCGACGACGGATTGCTGGCACCGGATTAACGCTGG 300
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 428 GGACAITACCAACGCCAGCGTCTGGTGGCGGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG 487

Qy 301 CAGTCTCGGTCAGGTCCCGGATTATTGTCGATGTCGACCAGCGTACCCGTAAAAACAGGCGAA 360
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 488 CAGCCTCGGTCAGGTCCCGGATTATTGTGCGATGTCGACCAGCGTACCCGTAAAAACAGGCGAA 547

Qy 361 GCCTTCGGCCGCGCGGCGGCTGGGCGCACCTGACTGTGGCGAATAACCGTACATACCCAGGCAGNTACG 420
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 548 GCCTTCGGCCGCGCGGCGGCTGGGCACCTGACTGTGTCCAATAACCGTACATACCCAGGCAGNTACG 607

Qy 421 TCCAGTCCCGATCCGTTGTTATTTAATCCTCTAAAAAACTGGGCGTTTGCCAACTGGATAACGCG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 608 TCCAGTCCCGATCCGTTATTTATTCCTTCTAAAAAACTGGGCGTTTGCCAACTGGATAACGCG 667

Qy 481 AACGTGACTGAGCGGATCCTCAGCAGGCGAGAGGTCAAATTTGCTAGCTTTACCGGGCAT 540
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 668 AACGTGACTGAGCGGATCCTCAGCAGGCGAGAGGTCAAATTGCTAGCTTTACCGGGCAT 727

Qy 541 GCCTTCGGCCGGAACTGGAACGGGGTGCTTAATTTCCGCAATCAAACTTGTGC 600
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 728 CGGCAAACGGGCGTTTCGCGGAACTGGAACGGGGTGCTTAATTTTCCGCAATCAAACTTGAAC 787

Qy 601 CTTAAACGTGAGAAACAGGACGAAAAGTGTTCATTAACGCAGGCATTACCATCGGAACTC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 788 CTTAAACGTGAGAAAACAGAATGAAAGCTGTAACTAACGCAGGCATTACCATCGGAACTC 847

Qy 661 AAGGTGAGCGCGCGACAAATGTTCATTAACCGGTGCGGTAAGCCTCGGCATCAAATGCTGACG 720
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 848 AAGGTGAGCGCGCGACAAATGTTTCATTAACCGGTGCGGTAAGCCTCGGCATCAAATGCTGACG 907

RESULT 14
US-08-910-798-1
; Sequence 1, Application US/08910798
; Patent No. 5876997
; GENERAL INFORMATION:
; APPLICANT: KRETZ
; TITLE OF INVENTION: NOVEL PHYTASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,798
; FILING DATE: August 13, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HAILE, PH.D., LISA A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/029001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 1:

Qy 721 GAGATATTTCTCCTGCAACAAGCAACAGGAAGCAGGGGGGTGGGGAAGGATCACC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 908 GAAATATTTCTCCTGCAACAAGCACAGGAATGCCGAGCCCGGGGTGGGGAAGGATCACT 967

Qy 781 GATTCACCAGTGGAACACCTTGCTAAGTTGCATAAGCGGCCAATTTATTGCTACAA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 968 GATTCACCAGTGGAACACCTTGCTAAGTTGCATAAGCGGCCAATTTATTATTGCTACAA 1027

Qy 841 CGCACGCCAGGAGGTTGCCCGCAGCCGCGCCACCCCCGTTATTGCGATTTCAGTACTG 900
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1028 CCCACGCCAGAGGTTGCCCGCAGTCGGGCCACCCCCGTTATTGGATTGATCAAGACAGCG 1087

Qy 901 TTGACGCCCCATCCACCGCAAAAAACAGGCGTATGGTGTGACATTACCGGCTACTG 960
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1088 TTGACGCCCCATCCACCGCAAAAAACAGGCGTATGGTGTGACATTACCGGCTCGTCTG 1147

Qy 961 TTTATTGCCGGACACGGATACTAAATCTGGGCAAATCTCGGCGCGCCGCACTGGTCAACTGG 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1148 TTTATTGCCGGACACGGATACTAAATCTGGGCAAATCTCGGCGCGCCGCACTGGTCAACTGG 1207

Qy 1021 ACGCTTCCGGTCAGCCGGATAACACGCCGGATAACACGCCGGGTGAACTGGTGTTTGAAACGTGG 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1208 ACGCTTCCAGGTCAGCCGGATAACACGCCGGGTGAACTGGTGTTTGAAACGGGCTGG 1267

Qy 1081 CGTCGGCTAAGCGGTAACAGCCAGTGGATTCAGGTTTCGCTGGTTCTTCCAGAACTTTACAG 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1268 CGTCGGCTAAGCGGATAACAGCCAGTGGATTCAGGTTTCGCTGGTTCTTCCAGACTTTACAG 1327

Qy 1141 CAGATGCCTGATAAAACGCCGCTTGTCATTAAATAACGGGCGGCCCCGGAGAGGTGAAACTGACC 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1328 CAGATGCCTGATAAAACGCCGCTTGTCATTAAATAACGGGCGGCCCCGGAGAGGTGAAACTGACC 1387

Qy 1201 CTGGCACGGATGTGAAGAGCGAAATGCCGAGCCATGTTGTTCGTTCGTTGGCAGGTTTTACGGCAA 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1388 CTGGCACGGATGTGAAGAGCGAAATGCCGAGCCATGTTGTTCGTTCGTTGGCGCGGTTTTACGGCAA 1447

Qy 1261 ATCGTGAATGAAGCACGCCATACCGGCGGCTGCAGTTTG 1296
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1448 ATCGTGAATGAAGCGCGCCATACCGGCGGCTGCAGTTTG 1483

```
;
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1272 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   IMMEDIATE SOURCE:
;     CLONE:
;   FEATURE:
;     NAME/KEY: PHYTASE
;     LOCATION:
; US-08-910-798-1

Query Match             91.5%;  Score 1210;  DB 2;  Length 1272;
Best Local Similarity   96.1%;
Matches 1272;  Conservative   0;  Pred. No. 0;  Mismatches  0;  Indels  51;  Gaps  1;

Qy    1 ATGAAAGCGATCTTAATCCCATTTTTATCTTCTGATTCCTGA                       60
        ||||||| |||||||||||||||||||||| |||||||||||
Db    1 ATGAAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCC                       60

Qy   61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGGAAAAGTGGAGT                      120
        ||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGTGGAAAGTGTCAGTTCATGT              120

Qy  121 GTGCGGTGCTTCAACCAAGGCCACCGCAACTGATGCAGGAGTGCAGGCCA              180
        ||||||||||||||||||||||||||||||||||||||||||
Db  121 GTGCGGTGCTTCAACCAAGGCCACCGCAACTGATGCAGGATGCAGCGCCATGGCCA        180

Qy  181 ACCTGGCCCGGTAAAACTGGGTTGGCTGACACCCGGCGCTATCTC                   240
        ||||||||||||||||||||||||||||||||||||||||||
Db  181 ACCTGGCCCGGTAAAACTGGGTTGGCTGACACCCGGCGCTATCTC                   240

Qy  241 GGACATTACCAACGCCACGCGTCTCGGTAGCCGACCGGATTCGTCTGGGGAAAAGGGGCTGCCCG  300
        ||||||||||||||||||||||||||||||||||||||||||
Db  241 GGACATTACCAACGCCACGCGTCTCGGTAGCCGACCGGATTCGTCTGGGGAAAAGGGGCTGCCCG  300

Qy  301 CAGTCTCTTGGTCAGGTTCGGCGGATTATTGCTGATGTCGGACGACGTACCCGGTAAAACAGG   360
        ||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGTCTCGTCAGGTCGGCGGATTATTGCTGATGTCGGACGACGTACCCGGTAAAACAGG      356

Qy  357 -----------------------------------------CCAGGCAGGATACG          369
                                                 |||||||||||||
Db  301 CAGTCTCGTCAGGTCGGCGGATTATTGCTGATGTCGGACGACGTACCCGGTAAAACAGG      356

Qy  361 GCCTTCGCGCGGCGGCTGCACCTGACGTGCAATAACCGTACATACCCAGGCAGAGATACG    420
        |||||||||||||
Db  357 -----------------------------------------CCAGGCAGGATACG        369

Qy  370 TCCAGTCCGATCCGTTATTTAATTCCTCTAAAAAACTGGACGTTGCCAACTGGATAACGCG    429
        ||||||||||||||||||||||||||||||||||||||||||
Db  370 TCCAGTCCGATCCGTTATTTAATTCCTCTAAAAAACTGGACGTTGCCAACTGGATAACGCG    429

Qy  421 TCCAGTCCGATCCGTTATTTAATCCTCTAAAAAACTGGACGTTGCCAACTGGATAACGCG     480
        ||||||||||||||||||||||||||||||||||||||||||
Db  421 TCCAGTCCGATCCGTTATTTAATCCTCTAAAAAACTGGACGTTGCCAACTGGATAACGCG     480

Qy  481 AACGTGACTGAGCGGATCCTCAGCGAGGGCGAGGAGGTCAATTGCTGACTTTACCGGGCAT    540
        ||||||||||||||||||||||||||||||||||||||||||
Db  481 AACGTGACTGAGCGGATCCTCAGCGAGGGCGAGGAGGTCAATTGCTGACTTTACCGGGCAT    540

Qy  430 AACGTGACTGAGCGGATCCTCAGCGAGGGCGAGGAGGTCAATTGCTGACTTTACCGGGCAT    489
        ||||||||||||||||||||||||||||||||||||||||||
Db  430 AACGTGACTGAGCGGATCCTCAGCGAGGGCGAGGAGGTCAATTGCTGACTTTACCGGGCAT    489

Qy  541 CGGCAAACGGGCGTTTCGGCGAACGGGGTCTTAATTTTCCGCAATCAAACTTGTGC        600
        ||||||||||||||||||||||||||||||||||||||||||
Db  490 CGGCAAACGGGCGTTTCGGCGAACGGGGTCTTAATTTTCCGCAATCAAACTTGTGC        549

Qy  601 CTTAAAACGTGAGAACAGGACGAGAAAGCTGTTCATTAACGGCAGCGCATTACCACTCGGGACTC  660
        ||||||||||||||||||||||||||||||||||||||||||
Db  550 CTTAAAACGTGAGAACAGGACGAGAAAGCTGTTCATTAACGGCAGCGCATTACCACTCGGGACTC  609

Qy  661 AAGGTGAGCGGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCGATCCAATGCTGACG    720
        ||||||||||||||||||||||||||||||||||||||||||
Db  610 AAGGTGAGCGGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCGATCCAATGCTGACG    669

Qy  721 GAGATATTTCTCTGCAACAAGCACAGGGAATCCCGGAGCCGGGGTGGGGAAGGATCACC         780
        ||||||||||||||||||||||||||||||||||||||||||
Db  670 GAGATATTTCTCTGCAACAAGCACAGGGAATCCCGGAGCCGGGGTGGGGAAGGATCACC         729

Qy  781 GATTCACACCAGTGGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTGCTACAA       840
        ||||||||||||||||||||||||||||||||||||||||||
Db  730 GATTCACACCAGTGGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTGCTACAA       789
```

```
Qy  841 CGCACGCCCAGAGGAGGTTGCCCCGCCCAGCCCCCGGTTATTGGATTTGATCATGGCAGCG        900
        ||||||||||||||||||||||||||||||||||||||||||
Db  790 CGCACGGCGCAGAGGTTGCCCCGCCCAGCCCCCGGTTATTGGATTTGATCATGGCAGCG          849

Qy  901 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGTGTGACATTCAGTCAGTGGCAGCG           960
        ||||||||||||||||||||||||||||||||||||||||||
Db  850 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGTGTGACATTCAGTCAGTGGCAGCG           909

Qy  961 TTTATTGCCGGACGACGGATACTAATCTGGCAAATCTCGGCGGCCGCACTCAACTGG           1020
        ||||||||||||||||||||||||||||||||||||||||||
Db  910 TTTATTGCCGGACGACGGATACTAATCTGGCAAATCTCGGCGGCCGCACTCAACTGG           969

Qy 1021 ACGCTTCCCGGTCAGCCGGATAACACGCCCCAGGTGGGTGAACTGGGTTGTTGAAACGCTGG      1080
        ||||||||||||||||||||||||||||||||||||||||||
Db  970 ACGCTTCCCGGTCAGCCGGATAACACGCCCCAGGTGGGTGAACTGGGTTGTTGAAACGCTGG      1029

Qy 1081 CGTCCGGCTAAGCGATAACAGCCACTGGATTCAGTTTCGCGTGGTCTTCAGAGACTTTACAG     1140
        ||||||||||||||||||||||||||||||||||||||||||
Db 1030 CGTCCGGCTAAGCGATAACAGCCACTGGATTCAGTTTCGCGTGGTCTTCAGAGACTTTACAG     1089

Qy 1141 CAGATGCGTGATAAAAACGCGCTGCTTCATTAAATAACGCGCCGCCGAGAGGTGAAACTGACC    1200
        ||||||||||||||||||||||||||||||||||||||||||
Db 1090 CAGATGCGTGATAAAAACGCGCTGCTTCATTAAATAACGCGCCGCCGAGAGGTGAAACTGACC    1149

Qy 1201 CTGGCAGGATGTGAAGAGCGAAAATCGCAGGGCATGTGTTCGTTGGTCAGGTTTTACGCAA     1260
        ||||||||||||||||||||||||||||||||||||||||||
Db 1150 CTGGCAGGATGTGAAGAGCGAAAATCGCAGGGCATGTGTTCGTTGGTCAGGTTTTACGCAA     1209

Qy 1261 ATCGTGAATGAAAGCACGCATACCGGCGTGCAGTTTGAGATCTCATCACCATCACCAATCAC    1320
        ||||||||||||||||||||||||||||||||||||||||||
Db 1210 ATCGTGAATGAAAGCACGCATACCGGCGTGCAGTTTGAGATCTCATCACCATCACCAATCAC    1269

Qy 1321 TAA                                                               1323
        |||
Db 1270 TAA                                                               1272
```

```
RESULT 15
US-09-489-039A-341
; Sequence 341, Application US/0948903A
; Patent No. 6610836
; GENERAL INFORMATION:
;   APPLICANT: Gary Breton et. al
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;   TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;   FILE REFERENCE: 2709.2004001
;   CURRENT APPLICATION NUMBER: US/09/489,039A
;   CURRENT FILING DATE: 2000-01-27
;   PRIOR APPLICATION NUMBER: US 60/117,747
;   PRIOR FILING DATE: 1999-01-29
;   NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 341
;   LENGTH: 1266
;   TYPE: DNA
;   ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-341

Query Match              4.8%;  Score 63.6;  DB 3;  Length 1266;
Best Local Similarity   48.7%;
Matches 202;  Conservative   0;  Pred. No. 1.7e-10;  Mismatches 210;  Indels   3;  Gaps  1;

Qy   86 AGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGTGTGCGTGCTCC---AACCAAGGCCA           142
        ||||||||||| |||| |||| | ||||| |||| || || ||       | ||||| ||
Db  113 AGCTGCAGCAGGTGCTGATCATCAGATGACCGGTTCATCGGTGTGGTTGCATAACTACG           172

Qy  143 CGCAACTGATGCAGGATGTCACCCCAGACGCATGGCAATCGGCCGTAAAAACTGGGGTT          202
        |||| |||| |   |  |     ||  || |  || || |||  | | |  |||||| ||
Db  173 GCAGCGTCGTGGACAGTCCACGCCCAAGGCCTGGCCCAGTGCCAAGTTGCCGCGCGCC          232

Qy  203 GGCTGACGACCGCGGGTGGTGGTGAGCTAATGCCTATCTCGGGACATTACCAACGCCGTC       262
        |||| ||   |   || | || ||  | || | |||| |  |  |||  ||  | ||| |
Db  233 AGCTGACCACCCAAAGGGCGGTGTGCTGGAGGTGTATAGCGGCCCACTATATGCGGAATGC     292
```

```
Qy   263 TGGTAGCCGACGGATTGCTTGGCGAAAAAGGGCTTGCCCGCAGTCTGGTCAGGTCGGCGATTA 322
             |||                    |||||||||||||
Db   293 TGGCGCAGCAGAAGCTGGTGACCAGCGGCGACGTGCCCGCGAAAACGCGGTTTATGCCT 352

Qy   323 TTGCTCGANTCGACGAGCGGTACCCGTAAAACAGGGCGAAGCCTTCGCCGCCGGGGCTGGGCAC 382
             ||||||                   |||||||                   ||||
Db   353 ACGCTAACAGCCTGCAGCCGCACCGTCGGCCACCGCGAGTTCTTTATCACCGGCGCGGCGTTCC 412

Qy   383 CTGACTGTGCAATAACCGTACATACCCAGGGCAGATACGTCCAGTCCCGATCCGGTTATTTA 442
             |   ||||              |||             |   |||        |   ||||
Db   413 CGGGCTGCGGCATCCCGGGTGCAATCACCCAGCCTCAGATNGGGCACCCATNGGACCCGACCTTCA 472

Qy   443 ATCCTCTAAAAAACTGGCGTTTGCCAACTGGANTAAACCCGAACGTGACTGACCGCGAT 497
             |   ||||              |   |||||              |||||
Db   473 ACCCGGGTTATTACCGACGACTCCCCGCCGTTTCGCCGAGAAAAGCCGCTGCCAGGCGGAT 527
```

```
                                GenCore version 5.1.9
                                Copyright (c) 1993 - 2006  Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on:         June 14, 2006, 10:31:46 ; Search time 39.7719 seconds
                                                    (without alignments)
                                                    7741.348 Million cell updates/sec

Title:          US-10-601-319-9
Perfect score:  1308
Sequence:       1 atgaaagcgatcttaatccc.........gcagttgagatctcaatcta 1308
Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0
Searched:       290227 seqs, 117694381 residues
Total number of hits satisfying chosen parameters:        580454

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      Published Applications NA New:*
                1: /EMC_Celera_SID3/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                2: /EMC_Celera_SID3/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                3: /EMC_Celera_SID3/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                4: /EMC_Celera_SID3/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                5: /EMC_Celera_SID3/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                6: /EMC_Celera_SID3/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                7: /EMC_Celera_SID3/ptodata/2/pubpna/US11_NEW_PUB.seq:*
                8: /EMC_Celera_SID3/ptodata/2/pubpna/US60_NEW_PUB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	37.6	2.9	1627	6	US-10-449-902-6979	Sequence 6979, Appl
2	34.8	2.7	1257	7	US-11-217-529-79920	Sequence 79920, A
3	34	2.6	2879	7	US-11-293-697-2180	Sequence 2180, Ap
4	33.4	2.5	2271	6	US-10-953-349-11004	Sequence 11004, Ap
5	32.4	2.5	1393	6	US-10-953-349-27515	Sequence 27515, A
6	32	2.4	1396	6	US-10-449-902-8987	Sequence 8987, Ap
7	31.8	2.4	2634	6	US-10-449-902-13868	Sequence 13868, Ap
8	31.4	2.4	3129	6	US-10-449-902-13863	Sequence 13863, A
9	31	2.4	1344	6	US-10-449-902-10760	Sequence 10760, A
10	31	2.4	2791	7	US-11-293-697-707	Sequence 707, App
11	30.8	2.4	3553	6	US-10-449-902-15013	Sequence 15013, A
12	30.6	2.3	3050	6	US-10-517-441-65	Sequence 65, App
13	30.2	2.3	422	6	US-10-449-902-965	Sequence 965, App
14	30.2	2.3	772	6	US-11-289-102-18030	Sequence 18030, Appl
15	30	2.3	3452	7	US-11-289-102-50	Sequence 50, Appl
16	30	2.3	3614	6	US-10-449-902-12893	Sequence 12893, A
17	29.8	2.3	1233	6	US-10-449-902-4509	Sequence 4509, Ap
18	29.8	2.3	1725	6	US-10-953-349-13409	Sequence 13409, A
19	29.8	2.3	2306	6	US-10-449-902-25657	Sequence 25657, A
20	29.8	2.3	2520	7	US-11-217-529-80383	Sequence 80383, A
21	29.8	2.3	2826	6	US-10-449-902-12802	Sequence 12802, A
22	29.6	2.3	1051	6	US-10-449-902-2911	Sequence 2911, Ap
23	29.6	2.3	1461	6	US-10-449-902-11476	Sequence 11476, A
24	29.6	2.3	2036	6	US-10-449-902-22962	Sequence 22962, A
25	29.6	2.3	15304	6	US-10-505-928-828	Sequence 828, App
26	29.4	2.2	364	7	US-11-154-977-74	Sequence 74, Appl
27	29.4	2.2	702	7	US-10-449-902-5189	Sequence 5189, Ap
28	29.4	2.2	718	7	US-11-154-977-68	Sequence 68, Appl
29	29.4	2.2	807	7	US-11-154-977-42	Sequence 42, Appl
30	29.4	2.2	1161	7	US-11-154-977-36	Sequence 36, Appl
31	29.4	2.2	1309	6	US-10-449-902-20908	Sequence 20908, Appl
32	29.4	2.2	1489	6	US-10-449-902-6188	Sequence 6188, Ap
33	29.4	2.2	1600	6	US-10-449-902-17204	Sequence 17204, A
34	29.4	2.2	1946	6	US-10-449-902-20839	Sequence 20839, A
35	29.4	2.2	2071	6	US-10-449-902-24250	Sequence 24250, A
36	29.4	2.2	2108	6	US-10-449-902-21152	Sequence 21152, A
37	29.4	2.2	2367	7	US-11-121-154-48	Sequence 48, Appl
38	29.4	2.2	2754	7	US-11-217-529-5697	Sequence 5697, Ap
39	29.4	2.2	16032	6	US-10-527-552-5	Sequence 5, Appli
40	29.2	2.2	2410	7	US-11-293-697-115	Sequence 115, App
41	29	2.2	3168	6	US-10-449-902-12853	Sequence 12853, A
42	29	2.2	3926	6	US-10-449-902-12996	Sequence 12996, A
43	28.8	2.2	956	6	US-10-449-902-11182	Sequence 11182, A
44	28.8	2.2	3457	7	US-11-293-697-1224	Sequence 1224, Ap
45	28.8	2.2	5862	6	US-10-505-928-25	Sequence 25, Appl

ALIGNMENTS

```
RESULT 1
US-10-449-902-6979/c
; Sequence 6979, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT:  National Institute of Agrobiological Sciences.
; APPLICANT:  Bio-oriented Technology Research Advancement Institution.
; APPLICANT:  The Institute of Physical and Chemical Research.
; APPLICANT:  Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE:  2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE:  2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE:  2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6979
;   LENGTH: 1627
;   TYPE: DNA
;   ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
;   DATABASE ACCESSION NUMBER: AK064110
;   DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-6979

Query Match             2.9%;  Score 37.6;  DB 6;  Length 1627;
Best Local Similarity   50.6%;
Matches    91;  Conservative    0;  Mismatches  89;  Indels    0;  Gaps    0;

Qy      115 CATGGTGTGCGTGCTCCACCAAGGCCACGCAACTGATGCAGGANTGTCACCCAGACGCA 174
            ||| |||||||| ||| ||||||  ||| ||||| || ||||| ||| |||||||||||
Db      645 CATCTTGCAGACGGTCGCCTCAACGAACCGCACCAGTCGAAGGTTCGTCGTCCCGGTCGAC 586

Qy      175 TGGCCAACCTGGCCGGGTAAAAACTGGGGTGAGCTGACACCGCGCCGGGTGGGTGAGCTAA 234
            ||||||| |||||| ||||    ||| || || || ||||| |||||| ||| |||| ||
Db      585 TTCGGGCGGTCGAACGGCGCCCTCTTCCTCGGAGGATGGCGCGCTGCGGGCGCTGAGCGCC 526

Qy      235 TATCTCGGACATTACTGCGGTCAGCGTTCTGGTTAGCCGACGGATTGCTGCCTAAATGTGC 294
            ||| ||| || || || |||||| |||| ||| |||| |||| |||| || |||  |||
Db      525 CATCACGGTGAATCCTCGTGTCCACCCTCTGCACGGCAGCAGATCGTACACGCTTATCCAAGGC 466
```

RESULT 2
US-11-217-529-79920/c
; Sequence 79920, Application US/11217529

; Publication No. US2006009612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 79920
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-79920

Query Match   65.4%; Score 34.8; DB 7; Length 1257;
Best Local Similarity   65.4%;
Matches   51; Conservative   2.7%; Pred. No. 0.061; Mismatches   27; Indels   0; Gaps   0;

Qy   671 CCGACTGTGTCTCATTANACCGTGCGTAAGCCTCCGCATCAANTGCTCACGAGAGATATTTC 730
Db   253 CCGAGTCTGTCTCAAAAATGGGTTTGGGTGAGCGGTGCCATNCAGCGCTGAAGAGAGCCGTATT 194

Qy   731 TCCTGCAACAAGCACACAGG 748
Db   193 TCATCCAAATCCCTAGAGG 176

RESULT 3
US-11-293-697-2180/c
; Sequence 2180, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2180
; LENGTH: 2879
; TYPE: DNA
; ORGANISM: Homo sapiens

Query Match   50.6%; Score 34; DB 7; Length 2879;
Best Local Similarity   50.6%;
Matches   82; Conservative   0; Pred. No. 0.19; Mismatches   80; Indels   0; Gaps   0;

Qy   134 CCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGACGCAACCTGGCCGGTAA 193
Db   388 CCGCGGCCTTGCTTATTGAGGAAGGAGGTGTCCCCAAAGGGGTTCCGGCCCCGCCAAGT 329

Qy   194 AACTGGGTGAGCTGACACCGCGCGGCGGTGGTGAGCTAATCGCCTTANTCTGGGACANTTACTGGC 253
Db   328 GCATGGTGTGCGGGGACTCGCCCAGCGCGGCGCTGATCATCTCGGCCGTGTGAGGTCCCGCTC 269

Qy   254 GTCAGCGTCTCGGTTAGCCGCGACGGANTTGCTGCCTAAANTGTGGGCT 295
Db   268 GGGAACGGGCGCTTGGGAGTGCCACCGAGCTGGGACACCAGTTGCT 227

RESULT 4

US-10-953-349-11004/c
; Sequence 11004, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCONDED THEREBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 11004
; LENGTH: 2271
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-11004

Query Match   60.4%; Score 33.4; DB 6; Length 2271;
Best Local Similarity   60.4%;
Matches   55; Conservative   0; Pred. No. 0.27; Mismatches   36; Indels   0; Gaps   0;

Qy   1212 TGAAGAGCGAAATGCGAGGGGCATGGTGTTCGGTTGGCAGGTTTTACGCAAATCCTGAATGA 1271
Db   1734 TGAACCGGGACATGTTCGGGGTGATATGTACCTGAGAGAGGGTTTCTGCAGATCTTGAACGA 1675

Qy   1272 AGCACGCATACCGCGTGCAGTTTGAGANTCT 1302
Db   1674 CCTGCGAAAACTAGTTTCCTTCTTTGANCTTT 1644

RESULT 5
US-10-953-349-27515/c
; Sequence 27515, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCONDED THEREBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 27515
; LENGTH: 1393
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-953-349-27515

Query Match   48.4%; Score 32.4; DB 6; Length 1393;
Best Local Similarity   48.4%;
Matches   90; Conservative   0; Pred. No. 0.46; Mismatches   96; Indels   0; Gaps   0;

Qy   603 TAAACGTGAGAAAACAGGAGCAAAAGCTGTTCATTAAACGCAGGCATTACCATCGGAAACTCAA 662
Db   1167 TAACCTTTAGGTACATCTTGCACGTTCTACCAACAGGTAAGCTTAAACAGCTGACAGTCCA 1108

Qy   663 GGTGAGCGCCGACTGGTGTCTCATTAAACCGGTGCGGTAAGCCTCCGCATCAANTGCTGACGGA 722
Db   1107 GGGGTGGGGCGCCGTGGGGGGTGAGTGGACTGACAACAGCTGACAAGTCACTATTTCACGTGTCCCCGG 1048

Qy   723 GATATTTCTCCTCGAACAAGCACAGGGAATGCCCGAGACGCCGGGGTGGGGGAGGATCACCGA 782
Db   1047 GATTAAGTCACTACACATACAAACGGTCCGGTGCTTAAGACAGGAAAGGAGGTCTCACAGA 988

Qy   783 TTCACA 788
Db   987 TTTACA 982

RESULT 6
US-10-449-902-8987

Sequence 8987, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A02O5Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 5791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8987
; LENGTH: 1396
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK106335
; DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-8987

Query Match   58.3%;   Score 32;   DB 6;   Length 1396;
Best Local Similarity   58.3%;   Pred. No. 0.63;
Matches   56;   Conservative   0;   Mismatches   40;   Indels   0;   Gaps   0;

Qy   679 GTTCTCATTAACCGGTGCGGTAAGCTCGGCATCAAATGCTGACGGAGATATTTCTCCTGCAA 738
          |||||||| |||||| |||||||||| | || ||| | | ||||| |||
Db   456 GTTCTCGTTCTCGGTCGGTCGGCCTTCTCGCCGAAGCGGGAGAGCTCGTACTTCCGAGAA 515

Qy   739 CAAGCACAGGGAATGCCGGAGCCGGGGTGGGGAAGG 774
          || | || |||| |||||||| |||||||| |||
Db   516 CCGGCCATGGGGGGGACAGGCGGCGCGGGGTGTGGCGG 551

RESULT 7
US-10-449-902-13868
Sequence 13868, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A02O5Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 5791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 13868
; LENGTH: 2634
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK111216
; DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-13868

Query Match   46.9%;   Score 31.8;   DB 6;   Length 2634;
Best Local Similarity   46.9%;   Pred. No. 1.1;
Matches   99;   Conservative   0;   Mismatches   112;   Indels   0;   Gaps   0;

Qy   199 GGTGGAGCTTGACCACCGCGCGCGGGTGGTGAGCTAATGGCCTATCTCCGACACATTACTGCGGTCAG 258
          |||| | |||||| ||| | |||| |||| |||| || | || |||||||| || |||

RESULT 8
US-10-449-902-13863/c
Sequence 13863, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A02O5Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 5791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 13863
; LENGTH: 3129
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK111211
; DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-13863

Query Match   47.7%;   Score 31.4;   DB 6;   Length 3129;
Best Local Similarity   47.7%;   Pred. No. 1.7;
Matches   92;   Conservative   0;   Mismatches   101;   Indels   0;   Gaps   0;

Db   317 GGGGTTCGGGCTCCACGTCGGCGGGTGTCGGGAAGGGCCCGGGTCGGCGGCCCGTCGG 376
Qy   259 CGTCTGGGTAGCCGACGCGGATTGCTGCCTAAAATGTAGGTGGCTGCCCGGCGAGTCTTCAGGTCGCG 318
          | | | | |||| ||| || | | || || ||| ||| |||||| ||
Db   377 CTTCGGTGTGGGGCCGCACGGTGGTCATCCTCGGGCGGGTTCGCCGCCGGACCTCGACCGCCCGGGA 436

Qy   319 ATTATTGCTGGATGTCGACGGAGCGTACCCGTAAAACAGCGGGAGAGCCTTCGCCGCCGGCTG 378
          || | || ||| | |||| ||| ||| || | | || ||||| ||| |||
Db   437 CTTCTGGTCGTCACCGCCATCCTTCCTTCGTCGAGGGGGGTGTTCAGCCGGGAGCA 496

Qy   379 GCACCTGACTGTGCAATAACCGTACATACCC 409
          || | |||||| || |||| ||| |||||
Db   497 CGAGCTCGATCTCGCAGGAGCAGCCGATGCAC 527

RESULT 9
US-10-449-902-10760/c
Sequence 10760, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.

Qy   999 CCGGCGCACTGGAGCTCAAACTGGACGCCTTCCCCGGTCAGCGGGATAAACACGCCGCCAGGTGG 1058
          | |||| | || |||| || | | ||||||| |||||| || |||||| || |||||||
Db   2942 CGAAGTTCCGGAGCTCAGACTCAGCACGTCGCGCTGCCCATGGCTTCCGTTGCGGGATAGGGG 2883

Qy   1059 TGAACTGGTGTTTGAACGCTGGGGTCGGGCTAAGCGGATAACAGCGCCAGTGGATTCAGGTTTC 1118
          || ||||| | | |||| ||||| | || | || ||| || ||| || |||||
Db   2882 TGGGTGGAACGGGTTACCGAGTTGGAGCCACAGAGGGTGAGTGCGGTCGGTCGGAAGCCATC 2823

Qy   1119 GCTGGGTCTTCCAGACTTTACAGCAGATGCGGTAAAAACGCCCCTGTCATTAAAATACCCC 1178
          | ||||||| || || |||||| || | ||||| | | |||||| |||| ||||||
Db   2822 GGTGAAACCCGAGAGTAGTTGGAGGAGCATGGAGGAGGAACCGGCCCCATTGCCGGCCGAGGGA 2763

Qy   1179 GCCCGGGAGAGGTG 1191
          |||| || |||||
Db   2762 ACCAGGAGTGCTG 2750

; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A020SY1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-12-11
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10760
; LENGTH: 1344
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK108108
; DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-10760

Query Match 53.8%; Score 31; DB 6; Length 1344;
Best Local Similarity 2.4%;
Matches 64; Conservative 0; Mismatches 1.4; Indels 0; Gaps 0;

```
Qy   258 GCGTCTGGTAGCCGACGGATTGCTGCTCTAAATGTGGCTCGGTCTGGTTCAGGTCGC 317
             ||||||| ||| || |||||| |||  ||||||||   | ||| ||||| |||
Db   289 GCGTCGGGTAGTCGATCGAYCCGTACGGGCCCGCCCGGGTGATCAGGTTAGT 230

Qy   318 GATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGAGGGAAGACCTTCGCCGCCGGGGC 376
             |||||  ||  |||||| |||||||||||| ||||||||||| |  ||||| || ||
Db   229 TGTTGTTGCCCGTTGTCGACGAGGAGTGCGCCGAACAGAGAAGAAGGCCGCGCTCCGCGGGAC 171
```

RESULT 10
US-11-293-697-707/c
; Sequence 707, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 707
; LENGTH: 2791
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-707

Query Match 55.0%; Score 31; DB 7; Length 2791;
Best Local Similarity 2.4%;
Matches 61; Conservative 0; Mismatches 2.2; Indels 0; Gaps 0;

```
Qy    59 CATTCGCTCAGAGTGAGCCGGAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATG 118
             ||||| ||| | ||| ||||| |||| ||| ||||| |||||  ||||
Db  1419 CCTTCGCAGYGAGYGTTACAGCTCTTAAGGCGGTGYGTCTGGAGYTGGTTCAYTCCTCCTG 1360

Qy   119 GTGTGCGGGTGCTGCTCCAACCACGCCACCCACCCTCAGGAGGAGTGTCACCCCAG 169
             ||||  ||||||||    ||||
Db  1359 GTGGGTTCGTGGTGTCTCGCTGGGCCTCAGAAGTGAAGCTGCAGAATCTTCGGCGG 1309
```

RESULT 11
US-10-449-902-15013
; Sequence 15013, Application US/10449902

; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.

; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A020SY1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-12-11
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 15013
; LENGTH: 3553
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK065548
; DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-15013

Query Match 47.8%; Score 30.8; DB 6; Length 3553;
Best Local Similarity 2.4%;
Matches 89; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy   297 CCCGCAGTCTCGGTCAGGTCGCGGCGAATTATTGCTGATGTTCGACGACGCGTACCCGTAAAACAGG 356
             ||||| ||| ||||| |||| ||  |||  ||| | | ||  || | | |||| ||||
Db  2688 CTCCGGATTTTGGACTTGGCCAAGCTCTAYGATGTGAAAAGGAGGACCCATAYCAGCACGAAAAT 2747

Qy   357 CGAAGGCCTTCGCCGCCGGGCGTGGCGCACCTGACTGTGCGAATAAACCGTACATACCCAGCGACAGA 416
             ||||||||||||  |||||||  | ||  ||||||| | ||| || ||| |||||| |||
Db  2748 CGCTGGCACATCGGCTATCTGGGCACCCGAGTATGCAATGAGGGGCCATCTGACTGAAAA 2807

Qy   417 TACGTCCAGTCCCGATCCGTGTATTTAATCCTCTAAAAAACTGGCGTTGCCAAACTGGGATAA 476
             ||||||||  |||  || ||||||   || || |||| |||
Db  2808 GGCTGAYGTTTTCGCAYTTCGGAGTTGTTGCTCTAGAAAACAGTTGCTGCGGCCGATCAAAACAC 2867

Qy   477 CGCGAA 482
             ||||||
Db  2868 CGACAA 2873
```

RESULT 12
US-10-517-441-65/c
; Sequence 65, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: FOEKENS, John
; APPLICANT: HARBECK, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMRICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARX, Almuth
; APPLICANT: HOEFLER, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell
; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; CURRENT FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147

```
; SEQ ID NO 65
; LENGTH: 3050
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-517-441-65

  Query Match             55.0%;  Score 30.6;  DB 6;  Length 3050;
  Best Local Similarity
  Matches  60;  Conservative  0;  Mismatches  49;  Indels  0;  Gaps  0;

Qy  160 GTCACCCCAGGACGGCATGGCCGAACCTGGGCCGGTAAAACTGGGTGAGCTGACACCGGCCGGT 219
         |||||  ||||| |||||| ||| |||| |||| |||||| |||| ||||| |||||||| ||
Db  281 GCCGCTCCCGTGCCCTGAGCATCATCCGCGCGGGCGTCCGGGGGGCTGAGCGGCCCGCCGAGGC 222

Qy  220 GGTGAGCTAATCGCCTACTCCGGACATTACTGCGTCAGCGTCTCGGTAG 268
         |||||||||  |||||||| |||| |||||||| |||||||||||||||
Db  221 TGTGCGCTAAACCTCCATTCTTAGAGATCCAACCCGGCAGGAGCGGGGGAG 173
```

RESULT 13
US-10-449-902-965/c
; Sequence 965, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A020SY1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 965
; LENGTH: 422
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK059042
; DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-965

```
  Query Match             52.8%;  Score 30.2;  DB 6;  Length 422;
  Best Local Similarity
  Matches  65;  Conservative  0;  Mismatches  58;  Indels  0;  Gaps  0;

Qy  658 CTCAAGGTGAGGCGCCGACTGTGTCTCATTAACCGGTGCGGTGAGGCCTGCCATCAATGCGTG 717
         |||| |||||||  |  ||| |||| ||| ||||  || |||| |||| ||||||
Db  224 CTCACGGTGAGTGCTGTGTATCCGGGCCACCTGAGTGAAGCAAACCAGCCACCAAAAATG 165

Qy  718 ACGGAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGACCCGGGCGTGGGGAAGGATC 777
         || ||||| |  || ||| | | | |||| ||  | |||| || |||||| ||||||  ||
Db  164 ATGAGGAAGAGATCTTGAGGCCGGAAGCAGCCGCGGGAGCGTCCTCGACGGCCGAGCGGCAGGACG 105

Qy  778 ACC 780
         |||
Db  104 ACC 102
```

RESULT 14
US-10-449-902-18030
; Sequence 18030, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A020SY1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 18030
; LENGTH: 772
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK068443
; DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-18030

```
  Query Match             45.3%;  Score 30.2;  DB 6;  Length 772;
  Best Local Similarity
  Matches 110;  Conservative  0;  Mismatches 133;  Indels  0;  Gaps  0;

Qy  137 AGGCCACGCAAACTGAATGCAGGATGTCACCCCAGAGACGCGCATGGCCGCAACCTGGCCCGGTAAAAC 196
         ||||||| |||||| | | |||| || |||||| | | | ||| ||||| ||||  |||  ||||| ||
Db   50 AGGCGCCGTTCCGTGTCACCTGTTCAACGACGACGACGCGCGGCCCAGCGCGGCGGCGGGGAGGCG 109

Qy  197 TGGGTGAGCTGACACCGCGCGGGTGGGTGAGCTAATCGCCTATCTCGGACATTACTGGCGTC 256
         |||||| || || ||| |||| |||| |||||||
Db  110 CCGGGCGCGGGCGTGACGGGTGCGCGGGGGAGACGGGTGCGGGCGTCGGCGGGAGAGCTGGGCCC 169

Qy  257 AGGGTCTCGTGAGCCGCGACGGATTTCTCTGCCTAAAATGTGGCTGCCCGGCACGTCTGGTCAGGTCG 316
         ||  | ||| |||  |||  |  |||  |||  |||||  |||| |||  | |  || |  ||  || |
Db  170 GGCGGGTGGGGGCCCCCGGGATGGGGCGGGCGTCCCGACGGGAATCATACTGGTCGCGAGAGATCG 229

Qy  317 CGATTATTGCTGAATGTCGACGACGGTACCCGTAAAACAGGCGAAGCCTTCGCCGTGGCGGGC 376
         |||| || ||| | | ||||| |||  ||| | || | ||  ||| ||  |||  ||  |||
Db  230 CCCGCGCGGTCGACGAGGAGGGCCCGGTCCCGCGCGGGCGGGGCCCGCCGGCGGGGCGGGGCGGC 289

Qy  377 TGG 379
         |||
Db  290 AGG 292
```

RESULT 15
US-11-289-102-50/c
; Sequence 50, Application US/11289102
; Publication No. US20060121511A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Hyerim
; APPLICANT: Clark, Edwin
; APPLICANT: Shaw, Peter M.
; TITLE OF INVENTION: BIOMARKERS AND METHODS FOR DETERMINING SENSITIVITY TO
; TITLE OF INVENTION: MICROTUBULE-STABILIZING AGENTS
; FILE REFERENCE: 10038 NP
; CURRENT APPLICATION NUMBER: US/11/289,102
; CURRENT FILING DATE: 2005-11-29
; PRIOR APPLICATION NUMBER: US 60/631,993
; PRIOR FILING DATE: 2004-11-30
; NUMBER OF SEQ ID NOS: 395
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 50
; LENGTH: 3452
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-289-102-50

```
  Query Match             64.3%;  Score 30;  DB 7;  Length 3452;
  Best Local Similarity
  Matches  45;  Conservative  0;  Mismatches  25;  Indels  0;  Gaps  0;

Qy  997 GGCGGGCGCCACTCGAGCTTCAACTGGACGCGCTTCCCGGTCAGCCCGGATAACACGCGCCCAGGT 1056
         |||   || |  |  |  |  ||| || | ||  ||  |||  ||  ||  ||  ||  |||  |||
```

```
Db   417 GACGTTGCGCTTGAGCCCGTCTAGACCTCGCTGGCTGAGGGTGATGACATGCCGGCCAGGT 358

Qy  1057 GGTGAACTGG 1066
          || | |||||
Db   357 GGGGCACTGG 348
```

Search completed: June 14, 2006, 14:23:14
Job time : 39.7719 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

Run on: June 14, 2006, 10:01:41 ; Search time 6867.62 Seconds
                 (without alignments)
                 10650.337 Million cell updates/sec

Title: US-10-601-319-9
Perfect score: 1308
Sequence: 1 atgaaagcgatcttaatccc...........gcagtttgagatctccatcta 1308

Scoring table: IDENTITY_NUC
                 Gapop 10.0 ; Gapext 1.0

Searched: 48236798 seqs, 2795665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   EST:*
                 1: gb_est1:*
                 2: gb_est3:*
                 3: gb_est4:*
                 4: gb_est5:*
                 5: gb_est6:*
                 6: gb_htc:*
                 7: gb_est2:*
                 8: gb_est7:*
                 9: gb_est8:*
                10: gb_est9:*
                11: gb_gss1:*
                12: gb_gss2:*
                13: gb_gss3:*
                14: gb_gss4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
c 1	592	45.3	616	2	BJ074127	BJ074127 BJ074127
c 2	580.2	44.4	853	13	CL662734	CL662734 PRI0142b
c 3	531.4	40.6	746	4	CA093060	SCCCL200
c 4	524.2	40.1	529	3	BU713770	SJAABUG02
c 5	488.2	37.3	595	5	CF326092	JMT1-05-
c 6	483.2	36.9	895	8	CN762997	ID0AAA5DC
c 7	446.8	34.2	868	8	CN754382	ID0AAA12D
c 8	442.8	33.9	864	8	CN757565	ID0AAA1DF
c 9	373.4	28.5	393	7	BE520240	M1B12STM
c 10	370.2	28.3	383	7	AW036132	EST274508
c 11	370.2	28.3	383	7	AW036134	AW036132
c 12	348.6	26.7	531	3	BU713769	SJAABUG01
c 13	295.8	22.6	706	8	CN759004	ID0AAA24B
c 14	295.4	22.5	354	4	AA545747	HBMSF1B4-
c 15	254.4	19.4	707	2	BJ619443	BJ619443
c 16	198	15.1	320	7	BE436403	EST4074B1
c 17	193.6	14.8	208	2	BM409016	EST583343
c 18	193.6	14.8	208	2	BM412806	EST587133
c 19	189.2	14.5	546	2	BJ040796	BJ040796 BJ040796
c 20	160.8	12.3	657	2	BG457132	NF10C05P USDA-FP_I
c 21	159.8	12.2	179	7	BE461872	EST413291
c 22	157.4	12.0	1171	8	CV672454	RET7SJ_01
c 23	151.2	11.6	508	2	BJ030203	BJ030203 BJ030203
c 24	143.2	10.9	254	7	BE520241	M1B12XTM
c 25	100.2	7.7	769	8	CO386885	AGENCOURT
c 26	88.8	6.8	863	9	CV672453	RET7SJ_01
c 27	59.8	4.6	772	14	DX010757	KBrB0Ö1C0
c 28	59.4	4.5	821	13	CL665831	PRI014d_E
c 29	59.4	4.5	842	13	CL665745	PRI014c_E
c 30	54	4.1	470	9	DN201839	USDA-FP_I
c 31	52.6	4.0	1175	9	CX943192	LamjGestG
c 32	43.8	3.3	392	10	H60613	yr44d12.81
c 33	41.8	3.2	274	3	BQ153360	NF03SC07I
c 34	40.8	3.1	250	14	DX065699	KBrB07JJ0
c 35	39.6	3.0	909	4	CA117609	SCBFLR108
c 36	39.4	3.0	600	4	CB547796	AMGNNUC:C
c 37	39.2	3.0	553	7	AW390601	MR3-ST019
c 38	39	3.0	733	2	BI152280	602917706
c 39	38.8	3.0	1160	9	DR148985	49350064
c 40	37.8	2.9	939	12	CL078996	CH216-154
c 41	37.6	2.9	347	4	CB705089	AMGNNUC:M
c 42	37.6	2.9	402	3	BU671248	NISC_1r06
c 43	37.6	2.9	418	5	CK739626	USDA-FP_6
c 44	37.6	2.9	425	4	CB696209	AMGNNUC:N
c 45	37.6	2.9	499	3	BP490202	BP490202 BP490202

ALIGNMENTS

RESULT 1
BJ074127/c

LOCUS   BJ074127   616 bp   mRNA   linear   EST 29-SEP-2003
DEFINITION BJ074127 NIBB Mochii normalized Xenopus tailbud library Xenopus
            laevis cDNA clone XL090j06 5', mRNA sequence.
ACCESSION   BJ074127
VERSION   BJ074127.1 GI:17504316
KEYWORDS   EST.
SOURCE   Xenopus laevis (African clawed frog)
ORGANISM   Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Xenopus.
REFERENCE   1 (bases 1 to 616)
AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and
            Kohara,Y.
TITLE   Expressed genes in X. laevis embryo
JOURNAL   Unpublished (2001)
COMMENT   Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp
            The information of this clone is available through the following
            URL.
            http://xenopus.nibb.ac.jp.
FEATURES   Location/Qualifiers
     source   1..616
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8355"
                     /clone="XL090j06"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 25"
                     /clone_lib="NIBB Mochii normalized Xenopus tailbud
                     library"
ORIGIN
Query Match   45.3%; Score 592; DB 2; Length 616;
Best Local Similarity   97.6%; Pred. No. 4.9e-175;

RESULT 2

```
Matches 601; Conservative  0; Mismatches  15; Indels  0; Gaps  0;

Qy    264 GGTAGCGACGGATTGCTGCCTAAAATGTGGGCTGCCCAGTCGTCTGGGTCAGGTCGCGATTAT 323
          ||||||||||||||||||||| ||| |||||||||||||||||||||||||||||||||||||
Db    616 GGTAGCGACGGATTGCTGGGCGGAAAAAAGGGGTCGCCCAGTCGTCGGGTCGCGGATTAT 557

Qy    324 TGCGTCGATGTCGACGAGCGGTACCCGTAAAACAGGCGGCAAGGCCTTCGCGCGCTGGCACC 383
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    556 TGCGTCGATGTCGACGAGCGGTACCCGTAAAACAGGCGGAAGCCTTCGCGCGCTGGCACC 497

Qy    384 TGACTCGTGCAATAACGGTACATACCCAGGCGAGATACGGTCCCGGATCCGTTATTTAA 443
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    496 TGACTCGTGCAATAACCGGTACATACCCAGGCGAGATACGGTCCCGGATCCGTTATTTAA 437

Qy    444 TCCTTCTAAAAACTGGCGTTTGCCACTGGATAAACGCGAACGTGACTGACCGGATCCTCGA 503
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    436 TCCTCTAAAAAACTGGGGTTTGCCAACTGGATAAACCGGATAAACGCGGGATCCTTCAG 377

Qy    504 GAGGGCAGGAGGGGTCAAATTGCTGACTTTACGGGGCATTATCAAAACGGCGTTTTCGCGAACT 563
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    376 CAGGGCAGGAGGGGTCAAATTGCTGACTTTACCGGGCAAAACGGCGTTTCGCGAACT 317

Qy    564 GGAACGGGTGCTTAAATTTCCGGCAATCAAAACTTGTGCCTTAAACGGGAAACAGGACGA 623
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    316 GGAACGGGTGCTTAATTTCCGGCAATCAAAACTGTGTCGCCTTAAACGTGAGAAACAGGACGA 257

Qy    624 AAGCTGTTCATTAACGGCAGGCATTACCCATTCGGAACTTCAAGGTGAGGCCCGACTGTGTCTC 683
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    256 AAGCTGTTCATTAACGGCGATTAACCATCGGAACTCAAGGTGAGCGCCCGACAATGTCTC 197

Qy    684 ATTAACCGGTGCGGTAAGCCTCGGCACACAATCGTGACGAGATATTTTCTCCTGCACACAAGC 743
          |||||||||||||||||||||||||||||||||||||||||||||
Db    196 ATTAACCGGTGCGGTAAGCCTCGGCATCAATGCTGACGGAGATATTTCTCCTGCAACAGC 137

Qy    744 ACAGGGAAATGCCGGAGCGGGTGGGGAAGGATCACGCGGATCCACGGATCAGCAGACACCTT 803
          |||||||||||||||||||||||||||||||||||||||||||
Db    136 ACTGGGGAAATGGCGGACCGGGGGTGGGGAAGGATCACCAGTGGAACAGACACCTT 77

Qy    804 GCTAAGTTTGCATAACGCGCAAATTTGATTTGCTACAACGGCCACGAGAGGGGGTTGCCCGGCAG 863
          |||||||||||||||||||||||||||||||||||||||||||||
Db    76 GCTAAGTTTGCATAACGGGCAATTTATTTTGCTACAACGGACACGCCCAGAGAGGTTGCCCGGCAG 17

Qy    864 CCGGCGCCACCCCGTTA 879
          |||||||||||||||||
Db    16 CCGGCGCCACCCCGTTA 1
```

```
LOCUS       CL662734           853 bp    DNA    linear    GSS 09-JUL-2004
DEFINITION  PRI0142b_B11 - PRI0142b.B21 (853) Mixed stage fosmid library of P.
            pacificus var. California Pristionchus pacificus genomic, genomic
            survey sequence.
ACCESSION   CL662734
VERSION     CL662734.1  GI:50150877
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 853)
  AUTHORS   Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
  TITLE     AppaDB: an AceDB database for the nematode satellite organism
            Pristionchus pacificus
  JOURNAL   Nucleic Acids Res. 32 (1), D421-D422 (2004)
   PUBMED   14681447
  COMMENT   Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 0049707160131
            Fax: 0049707161498
            Email: ralf.sommer@tuebingen.mpg.de
```

This library was generated at Caltech, Pasadena, USA and end
sequenced at Vancouver, Canada.
Seq primer: T7
Class: fosmid ends.

```
FEATURES             Location/Qualifiers
     source          1..853
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Mixed stage fosmid library of P. pacificus
                     var. California"
                     /note="Vector: pEpifos-5 Fosmid vector"
ORIGIN
```

```
Query Match        44.4%;  Score 580.2;  DB 13;  Length 853;
Best Local Similarity  98.7%;  Pred. No. 2.8e-171;
Matches  585;  Conservative  0;  Mismatches  8;  Indels  0;  Gaps  0;

Qy    708 ATCAATGCTGACGGAGATATTTCTCCTGCAACAGCACAGGGAAATGCCGGAGCCCGGGGTG 767
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATCAATGCTGACGGAGATATTTCTCTGCAACAGCACAGGGAAATGCCGGACCCGGGGTG 60

Qy    768 GGGAAGGATCACCGATTCACCACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAAATT 827
          |||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGGAAGGATCACCGATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATT 120

Qy    828 TGATTTGCTACAACGCACGCCAGAGGTTGCCCGGCCAGCCGGCGCCACCCCGTTATTAGATTT 887
          ||||||||||||||||||||||||||||||||||||||||||
Db    121 TTATTTTGCTACAACGGACACGCCAGAGAGGTTGCCCGGCAGCCGGCGCCACCCCGTTATTGGATTT 180

Qy    888 GATCAAGACAGACGGTTGACGCCCATCCACCCGCAAAAAACAGGGGTTATGGTGTGACATTACC 947
          |||||||||||||||||||||||||||||||||||||||||||||
Db    181 GATCATGGACAGCGTTGACGCGCCCATCACCCGCAAAAACAGGGGTTATGGTGTGACATTACC 240

Qy    948 CACTTCAGTGTGCTGTTTATCGCCGGACACGGATAACAGCCAGTGCAAAATCTCGGGCGGCACT 1007
          ||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CACTTCAGTCGTACGTTTATTGCCGGAGCATTAACTCTAATCTGGGAAATCTCGGCGCGCAACT 300

Qy   1008 GGAGCTCAACTGGACACGGCTTCCCGGGTCAGCCGGGATAACAACGCCGCCGCCAGGTGGTGTGT 1067
          |||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGAGCTCAACTGGACACGGCTTCCCGGGTCAGCCGGGATAACAACGCCGCCCAGGTGGTGAACTGGT 360

Qy   1068 GTTTGAACGCTTGGGCGTCGGGCTAAGGCCACGTAACAGCCACTGGGATTCAGTTTCGCTGGGCACT 1127
          |||||||||||||||||||||||||||||||||||||||||||||
Db    361 GTTTGAACGCTTGGGCGTCGGGCTAAGGCCATTAACAGCCACTGGGATTCAGTTTCGCTGGGCTCTT 420
```

RESULT 3

```
LOCUS       CA093060           746 bp    mRNA    linear    EST 23-SEP-2003
DEFINITION  SCCCL2001D10.b CL2 Saccharum officinarum cDNA clone SCCCL2001D10
            3', mRNA sequence.
ACCESSION   CA093060
VERSION     CA093060.1  GI:34946367
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharinae; Saccharum; Saccharum officinarum
```

```
Qy   1128 CCAGACTTTACAGCAGATGGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCGGAGA 1187
          ||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CCAGACTTTACAGCGAGTGGCGTGATAAAAACGCCGCTGTCATTAAATTACGCCGCCGAGA 480

Qy   1188 GGTGAAACTGACCCTGGCAGGATGTGAAGAGTGGGCGAAAATGCGCAGGGAGCATGTGTTCGTTGGC 1247
          |||||||||||||||||||||||||||||||||||||||||||||
Db    481 GGTGAAACTGACCCTGGCAGGATGTGAAGAGGTGGCGAAATGCGCAGGGAGCATGTGTTCGTTGGC 540

Qy   1248 AGGTTTTTACGCAAATCCTGAATGAACGAAGGCACGCAATACCCGGCGGTGCCAGTTTGAGAT 1300
          |||||||||||||||||||||||||||||||||||||||||||
Db    541 AGGTTTTTACGCAAATCCTGAATGAAGGCACGCAATACCCGGCGGTGCCAGTTTGTTAAT 593
```

REFERENCE   1 (bases 1 to 746)
AUTHORS   Vettore,A.L., da Silva,P.R., Kemper,E.L. and Arruda,P.
TITLE   The libraries that made SUCEST
JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT   Contact: Arruda P
         Centro de Biologia Molecular e Engenharia Genetica
         Universidade Estadual de Campinas
         Caixa Postal 6010, 13083-970, Campinas SP, Brazil
         Tel: 55 19 3788 1137
         Fax: 55 19 3788 1089
         Email: parruda@unicamp.br
         Clone distribution: clone distribution information can be found
         through the Brazilian Clone Collection Center (BCCC) at
         http://www.bccenter.fcav.unesp.br
         Plate: 001 row: D column: 10
         Seq primer: M13/Forward primer.
FEATURES   Location/Qualifiers
     source   1..746
                /organism="Saccharum officinarum"
                /mol_type="mRNA"
                /db_xref="taxon:4547"
                /clone="SCCCL2001D10"
                /lab_host="XL1Blue MRF'."
                /clone_lib="CL2"
                /note="Organ: Pool of sugarcane calli submitted to low
                (4oC) and high (37 C) temperature stress; Vector:
                pBluescript; Site_1: EcoRI; Site_2: XhoI; An
                unidirectional cDNA library generated from [Pool of
                sugarcane calli submitted to low (4oC) and high (37 C)
                temperature stress]. cDNA was prepared from polyA+ mRNA
                using ZAP - cDNA Synthesis Kit (Stratagene). The
                double-strand cDNAs were fractionated in a sepharose CL-2B
                40cm-columns and fragments sizing between 0.8 and 1.5 Kb
                were directionally cloned into the vector. Details of
                each source of RNA and library construction can be
                obtained at http://sucest.lad.ic.unicamp.br/public"
ORIGIN
     Query Match
     Best Local Similarity   97.3%;
     Matches 549; Conservative   0; Mismatches   14; Indels   1; Gaps   1;

Qy   595 TTGTGCCTTAAACGTGAGAAACAGGACGGAAAGTGTTCATTAACGCAGGCATTACCATCG 654
           ||| ||||||||||||||||||||||||||||||||||||||||||||
Db   184 TTGCGTACTGTCAGTGAGAAACAGGACGGAAAGCTGTTCATTAACGCAGGCATTACCATCG 243

Qy   655 GAACTCAAGGTGAGCGGCCGACTGTGTCTTCATTAACCGTGCGTAAGCCTTCGGCATCAATG 714
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   244 GAACTCAAGGTGAGCGGCCGACAAATGTCTCATTAACCGTGCGGTAAGCCTCGCGCATCAATG 303

Qy   715 CTGGACGGAGATATTCTCTCGCAACAAGCACAGGAATGCCGGAGCCGGGGTGGGGAAGG 774
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   304 CTGGACGGAGATATTTCTCTCGCAACAAGCACAGGAATGCCGGAGCCCGGGGTGGGGAAGG 363

Qy   775 ATCACCGGATTCACACCAGTGGAACACCTTGCTAAGTTGCATAACGGCCAATTGATTTG 834
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   364 ATCACCGGATTCACACCAGTGGAACACCTTGCTAAGTTGCATAACGGCCAATTTTATTTG 423

Qy   835 CTTACACACGCACCGCCAGAGGTTGCCCGGCGCCGCGCCCCGTTATTGATTTGATCAAG 894
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   424 CTTACACACGCACCGCCAGAGGTTGCCCGAGCCGCGGTTATTGATTTGATCAAG 483

Qy   895 ACAGGCGTTGAGCGCCCCATCCACCGCAAAAACAGGGCGTPATGGTGTGACATTACCCACTTCA 954
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   484 ACAGGCGTTGACCGCCCATCCACCGCANAAACAGGCGTTATGGTGTGACATTACCACTTCA 543

Qy   955 GTGGCTGTTTATCGCCGGACACAGATACTPAATTCCTGGGACACAGGCACTGGGACGTC 1014
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   544 GTGCGTGTTTATCGCCGGACACAGATACTPAATCTGGACAAAATCCTCGGCACTGGGACGTC 603

Qy   1015 AACTGGACGCTTCCCGGTTCAGCCGGATAACACGCCCGGCCAGGGTGGTGGGTGAACTGGGTG 1074
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   604 AACTGGACGCTTCCCGGTTCCGGTAGCCCGGATAACACGCCCGGCCAGGGTGGTGGGTGAACTGGGTG-TTGAA 662
           ||| ||||||||||||||||||||||||||||||||||||||||||||||
Qy   1075 CGGCTGGGCGTCGGGCTAAAGCCGATTAACACAGCCAGTGGATTCAGGTTTTCCGCTGGGTCTTCCAGACT 1134
           ||||||||| |||||||||||||||||||||||||||||||||||||||||||| ||
Db   663 CGGCTGGGCGTCGTCNGCGTAAGCGGATAACACAGCCAGTGGGATTCAGGTTTCGGCTGGGTCTTCCAGACT 722

Qy   1135 TTACAGCACAGATGCCTGATAAAACG 1158
           ||||||||||||||||||||||||||
Db   723 TTACAGCANATGCCGTGATAAAACG 746

RESULT 4
LOCUS   BU713770   529 bp   mRNA   linear   EST 23-OCT-2003
DEFINITION   SJAABUG02 Adult SjC 7/94 Schistosoma japonicum cDNA similar to
           pdb|1DKP|A Chain A, Crystal Structure Of Phytate Complex Of
           Escherichia Coli Phytase At Ph 6.6. Phytate Is Bound With Its
           3-Phosphate In The Active Site; Hg2+ Cation Acts As An
           Intermolecular Bridge, mRNA sequence.
ACCESSION   BU713770
VERSION   BU713770.1   GI:28321126
KEYWORDS   EST.
SOURCE   Schistosoma japonicum
   ORGANISM   Schistosoma japonicum
           Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
           Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE   1 (bases 1 to 529)
AUTHORS   Hu,W., Yan,Q., Shen,D.K., Liu,F., Zhu,Z.D., Song,H.D., Xu,X.R.,
           Wang,Z.J., Rong,Y.P., Zeng,L.C., Wu,J., Zhang,X., Wang,J.J.,
           Xu,X.N., Wang,S.Y., Fu,G., Zhang,X.L., Wang,Z.Q., Brindley,P.J.,
           McManus,D.P., Xue,C.L., Feng,Z., Chen,Z. and Han,Z.G.
TITLE   Evolutionary and biomedical implications of a Schistosoma japonicum
           complementary DNA resource
JOURNAL   Nat. Genet. 35 (2), 139-147 (2003)
PUBMED   12973349
COMMENT   Contact: Zeguang Han
           Chinese National Human Genome Center at Shanghai
           351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
           201203, P. R. China
           Tel: 86-21-50801919(ex.45)
           Fax: 86-21-50801922
           Email: hanzg@chgc.sh.cn.
FEATURES   Location/Qualifiers
     source   1..529
                /organism="Schistosoma japonicum"
                /mol_type="mRNA"
                /strain="Chinese (Anhui) strain"
                /db_xref="taxon:6182"
                /sex="Male and female"
                /tissue_type="Whole body"
                /dev_stage="Adult worms"
                /lab_host="Mouse and rabbit"
                /clone_lib="Adult SjC 7/94"
                /note="Vector: Lambda ZAP-II XR.; Site_1: EcoR I; Site_2:
                Xho I; Several hundred adult Schistosoma japonicum
                (Anhui, P.R. China, strain), of mixed sex, were perfused
                from the mesenteries of experimentally infected mice and
                rabbits at the Queensland Institute of Medical Research,
                Brisbane, Australia (QIMR), and stored for several months
                in liquid nitrogen. Subsequently, mRNA was isolated at the
                QIMR from lysates of these worms by oligo dT
                chromatography, using a kit from Pharmacia. The mRNA was
                then shipped to Clontech, Palo Alto, CA, USA, who
                constructed a cDNA library. First strand synthesis was
                primed with an oligo-dT-XhoI-primer and synthesized using
                M-MLV reverse transcriptase. Second strand synthesis was
                accomplished with RNAse H and T4 DNA polymerase. The
                double stranded cDNA was ligated to EcoRI linkers,
                digested with EcoRI and XhoI, and ligated into the
                phagemid vector lambda ZAP II XR. After construction of
                this directional library by Clontech, it was returned to
                the QIMR. During analysis of the library at the QIMR, we
                have found that a small percentage, 2% to 3%, of the

COMMENT   clones contain inserts that appear to be highly homologous
           to sequences from salmonid fishes, as determined by
           homology comparisons using BLAST and by Southern
           hybridization analysis to genomic DNA from salmon (Sigma
           Chemical Co, St. Louis, MO) under stringent washing
           conditions. The remainder of the clones appear to contain
           S. japonicum sequences."

ORIGIN

Query Match   40.1%; Score 524.2; DB 3; Length 529;
Best Local Similarity 99.4%; Pred. No. 1.3e-153;
Matches 526; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy   670 GCCGACTGTGTCTCATTAACCGGTTGCGGTAAGCTTCGCATCAATGTCGATGACGGAGATATTT 729
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GCCGACAAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGTCGATGACGGAGATATTT 60

Qy   730 CTCCTGCAACAAGCACAGGGAATGCCCGAGCCCGGGGGGGTGGGGGAAGGAGATCACCCGATTCACAC 789
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTCCTGCAACAAGCACAGGGAATGCCCGAGCCCGGGGGGGTGGGGGAAGGAGATCACCCGATTCACAC 120

Qy   790 CAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTGATTTGCTACAACGCACGCCA 849
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTTATTTGCTACAACGCACGCCA 180

Qy   850 GAGGGTTGCCCGCCCAGCCGCCGCCGCCGCCGTTATTAGATTTGATCAAGACAGACGCGGTTGACGCCC 909
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GAGGGTTGCCCGCCCAGCCGCCGCCGCCGTTATTAGATTTGATCAAGACAGACGCGGTTGACGCCC 240

Qy   910 CATCCACCGCAAAAAACAGGGCGTATGGGTGTGACATTACCCACTTCAGTGCTGCGTTTATCGCC 969
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CATCCACCGCAAAAAACAGGGCGTATGGGTGTGACATTACCCACTTCAGTGCTGCGTTTATCGCC 300

Qy   970 GGACAGCGATACTAATCTTGGCAAAATCTCGGCGCGCGCACTGAGCTCAACTGGACGCTTCCC 1029
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GGACAGCGATACTAATCTTGGCAAAATCTCGGCGCGCGCACTGAGCTCAACTGGACGCTTCCC 360

Qy   1030 GGTCAGCCGGATAACACGCGCCCCCCAGGTGGTGTGAACTCGGTGTTTGAAGGCTGCGTGGCCGTTCGGCTA 1089
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGTCACCCGGATAACACGCGCCCCCCAGGTGGTGAACTCGGTGTTTGAAGGCTGCGTGGCCGTTCGGCTA 420

Qy   1090 AGCGATAACACGCCAGTGGATTCAAGGTTTCGCTGCGGTCCTTCAGACTTTACAGGAGATGCGGT 1149
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AGCGATAACACGCCAGTGGATTCAAGGTTTCGCTGGGTCCTTCCAGACTTTACAGGAGATGCGGT 480

Qy   1150 GATAAAAACGCGCCGCTGTCATTAAATACGCCGCCCGGAGAGGGTGAAAACTGA 1198
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GATAAAAACGCGCCGCTGTCATTAAATACGCCGCCCGGAGAGGGTGAAAACTGA 529

RESULT 5
CF226092/c
LOCUS   CF226092   595 bp   mRNA   linear   EST 18-AUG-2003
DEFINITION JMT1--05-B09.g1 AtJMT-overexpressing transgenic rice lambda phage
            cDNA library (JMT1) Oryza sativa (japonica cultivar-group) cDNA
            clone JMT1--05-B09, mRNA sequence.
ACCESSION   CF226092
VERSION   CF226092.1 GI:33800445
KEYWORDS   EST.
SOURCE   Oryza sativa (japonica cultivar-group)
  ORGANISM Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
            clade; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1 (bases 1 to 595)
AUTHORS   Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
            Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE   Large-scale Sequencing Analysis of Rice ESTs
JOURNAL   Unpublished (2003)
COMMENT   Contact: Nahm B.H.
            Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
            of Bioscience and Bioinformatics, Myongji University
            Yongin, Kyeonggi, Korea

FEATURES   Location/Qualifiers
     source   1..595
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nackdong"
                     /db_xref="taxon:39947"
                     /clone="JMT1--05-B09"
                     /tissue_type="leaf"
                     /dev_stage="14 days after germination"
                     /lab_host="E.coli SOLR"
                     /clone_lib="AtJMT-overexpressing transgenic rice lambda
                     phage cDNA library (JMT1)"
                     /note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
                     XhoI; cDNA was inserted into lamda Uni-ZAP XR vector at 5'
                     end with EcoRI and 3' end with XhoI site. mRNA was
                     prepared from Arabidopsis Jasmoinate Carboxyl
                     methyltransferase overexpression line."

ORIGIN

Query Match   37.3%; Score 488.2; DB 5; Length 595;
Best Local Similarity 99.4%; Pred. No. 3.1e-142;
Matches 490; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy   808 AGTTTGCATAACGCGCAATTTGATTTGCTACAACGCACGCCAGAGGGTTGCCCGCCCAGCCGC 867
          ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   595 AGTTTGCGATAACGCGCAATTTTATTTGCTACAACGCACGCCAGAGGGTTGCCCGCCCAGCCGC 536

Qy   868 GCCACCCCCGTTATTAGATTTGATCAAGACAGACGCGGTTGACGCCCCATCCACCGCAAAAAACAG 927
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   535 GCCACCCCCGTTATTAGATTTGATCAAGACAGACGCGGTTGACGCCCCATCCACCGCAAAAAACAG 476

Qy   928 GCGTATGGGTGTGACATTACCCACTTCAGTGCTGCGTTTATCGCCGGACACGATAACTAATCTG 987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   475 GCGTATGGGTGTGACATTACCCACTTCAGTGCTGCGTTTATCGCCGGACAGATAACTAATCTG 416

Qy   988 GCAAAATCTCGGCGGCGCGCACTGAGCTCAACTGGACGCTTCCCGGGTCAGCCGGATAACACG 1047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   415 GCAAAATCTCGGCGCGCGCACTGAGCTCAACTGGACGCTTCCCGGGTCAGCCGGATAACACG 356

Qy   1048 CCGCCCAGGTGGTGGTGAACTGGTGTTTGAAGCGCTGCGGCGTGGCCGTTAAGCCAGTGG 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   355 CCGCCCAGGTGGTGAACTGGTGTTTGAAGCGCTGCGGCGTGGCCGTTAAGCCAGTGG 296

Qy   1108 ATTCAGGTTTCGCTGGTCCTTCCAGACTTTACAGCAGATGCGGTGATAAAAACGCGCCGCTGTCA 1167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   295 ATTCAGGTTTCGCTGGTCTTCCAGACTTTACAGCAGATGCGGTGATAAAAACGCGCCGCTGTCA 236

Qy   1168 TTAAATACGCCGCCCGGAGAGGGTGAAAACTGACCCTGCGCAGGATGTGAAGAGCGAAATGCG 1227
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   235 TTAAATACGCCGCCCGGAGAGGGTGAAAACTGACCCTGGCCAGGATGTGAAGAGCGAAATGCG 176

Qy   1228 CAGGGGCATGTGTTCGTTGGCAGGTTTTACGCAAATGTGTGAATGAAGCACGCGATACCGGCG 1287
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   175 CAGGGGCATGTGTTCGTTGGCAGGTTTTACGCAAATCGTGAATGAAGCACGCGATACCGGCG 116

Qy   1288 TGCAGTTTGAGAT 1300
          ||||||||||
Db   115 TGCAGTTTGTAAT 103

RESULT 6
CN762997/c
LOCUS   CN762997   895 bp   mRNA   linear   EST 20-MAY-2004
DEFINITION ID0AAA5DC09RM1 ApMS Acyrthosiphon pisum cDNA clone ID0AAA5DC09 5',
            mRNA sequence.
ACCESSION   CN762997
VERSION   CN762997.1 GI:47536920
KEYWORDS   EST.
SOURCE   Acyrthosiphon pisum (pea aphid)
  ORGANISM Acyrthosiphon pisum

Left column

```
REFERENCE   1  (bases 1 to 895)
AUTHORS     Hunter,W., Martinez-Torres,D., Rahbe,Y., Sabater-Munoz,B.,
            Stern,D., Tagu,D. and Wincker,P.
TITLE       An expressed sequence tags database for the pea aphid Acyrthosiphon
            pisum
JOURNAL     Unpublished (2004)
COMMENT     Contact: D. Tagu
            INRA Rennes
            UMR BiO3P, BP 35327, F-35653 Le Rheu Cedex France
            Tel: +33.2.23.48.51.65
            Fax: +33.2.23.48.51.50
            Risk of contamination by bacterial sequences from obligatory
            (Buchnera) or facultative endosymbionts. These sequences were
            obtained in the frame of the International Consortium of Aphid
            Genomics in collaboration with Genoscope
            PCR Primers
            FORWARD: CAGGAAACAGCTATGACC
            Plate: 5   row: C   column: 9.
FEATURES             Location/Qualifiers
     source          1..895
                     /organism="Acyrthosiphon pisum"
                     /mol_type="mRNA"
                     /cultivar="developmentstage"
                     /db_xref="taxon:7029"
                     /clone="ID0AAA5DC09"
                     /tissue_type="whole insect"
                     /dev_stage="nymphs and adults"
                     /lab_host="Xl1-Blue"
                     /clone_lib="ApMS"
                     /note="Vector: pBS-SK minus; Site_1: EcoRI; Site_2: XhoI;
                     Sample name: ID0AAA ; Plant growth place: Department of
                     Ecology & Evolutionary Biology, Princeton University ;
                     Soil conditions: Soil ; Sowing date: 01/06/1999 ;
                     Harvesting date: 01/06/1999 ; Stress date: no stress ;
                     Description: Aphids inoculated on one-week old Vicia faba
                     under non-sterile conditions. All parthenogenetic stages
                     and both winged and wingless adults were collected for
                     library construction. ; experimental condition: long
                     photoperiod (16-hr light/8-hr dark at 18 c)"
ORIGIN

Query Match         36.9%;  Score 483.2; DB 8; Length 895;
Best Local Similarity   99.4%;
Matches    485;  Conservative   0;  Mismatches   3;  Indels   0;  Gaps   0;

Qy    813 GCATAACGCGCAATTGATTTGATTTGCTACAACGCACGCACGAGGGTTGCCCGCGAGCGCCAC 872
          ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db    895 GCATAACGCGCAATTTTATTTGCTACAACGCACGCACGAGGGTTGCCCGCGAGCGCCAC 836

Qy    873 CCGGTTATTAGATTTGATCAAGACAGCGGTTGACGCCCCATCCACCGCAAAAAACAGGCGTA 932
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    835 CCGGTTATTAGATTTGATCAAGACAGCGGTTGACGCCCCATCCACCGCAAAAAACAGGCGTA 776

Qy    933 TGGTGTGACATTACCCACTTCAGTGCTGTTTATCGCGGGACACGATACTAATCTGGGCAA 992
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    775 TGGTGTGACATTACCCACTTCAGTGCTGTTTATCGCGGGACACGATACTAATCTGGGCAA 716

Qy    993 TCTCGGCGGGGGCGCACTCGGAGCTCAAACTCGGAGCGGTTCCCGGTACAGCGCCGCCGCC 1052
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    715 TCTCGGCGGGGGCGCACTCGGAGCTCAAACTCGGAGCGGTTCCCGGTACAGCGCCGCCGCC 656

Qy   1053 AGGTGGTGAACTGGGTGTTGAAGGCTCGGCGTCGGCGTGGGCTAAGCGATAACAGCCAGTGGGATTCA 1112
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    655 AGGTGGTGAACTGGGTGTTGAAGGCTCGGCGTCGGCGTGGGCTAAGCGATAACAGCCAGTGGATTCA 596

Qy   1113 GGTTTCCGTCGGTCTTCAGCAGGATGCGTGGGCTGCGGCTCGGCTAAGCGTGTCATTTAA 1172
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    595 GGTTTCCGTCGGTCTTCAGCAGTTTACAGCAGATGCGTGGGCTGCGGCTCGGCTAAGCGTGTCATTTAA 536

Qy   1173 TACGCGCGCCCGGAGAGGTGAAACTGACCCTGGCAGGATGTGAAGAGCGGAAATGCGCAGG 1232
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    535 TACGCGCGCCCGGAGAGGTGAAACTGACCCTGGCAGGATGTGAAGAGCGGAAATGCGCAGG 476
```

Right column

```
RESULT 7
LOCUS       CN754382/c               868 bp    mRNA            linear
DEFINITION  ID0AAA12BD01RM1 ApMS Acyrthosiphon pisum cDNA clone ID0AAA12DE01
            5', mRNA sequence.
ACCESSION   CN754382
VERSION     CN754382.1  GI:47519379
KEYWORDS    EST.
SOURCE      Acyrthosiphon pisum (pea aphid)
ORGANISM    Acyrthosiphon pisum
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
            Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
REFERENCE   1  (bases 1 to 868)
AUTHORS     Hunter,W., Martinez-Torres,D., Rahbe,Y., Sabater-Munoz,B.,
            Stern,D., Tagu,D. and Wincker,P.
TITLE       An expressed sequence tags database for the pea aphid Acyrthosiphon
            pisum
JOURNAL     Unpublished (2004)
COMMENT     Contact: D. Tagu
            INRA Rennes
            UMR BiO3P, BP 35327, F-35653 Le Rheu Cedex France
            Tel: +33.2.23.48.51.65
            Fax: +33.2.23.48.51.50
            Risk of contamination by bacterial sequences from obligatory
            (Buchnera) or facultative endosymbionts. These sequences were
            obtained in the frame of the International Consortium of Aphid
            Genomics in collaboration with Genoscope
            PCR Primers
            FORWARD: CAGGAAACAGCTATGACC
            Plate: 12   row: E   column: 1.
FEATURES             Location/Qualifiers
     source          1..868
                     /organism="Acyrthosiphon pisum"
                     /mol_type="mRNA"
                     /cultivar="developmentstage"
                     /db_xref="taxon:7029"
                     /clone="ID0AAA12DE01"
                     /tissue_type="whole insect"
                     /dev_stage="nymphs and adults"
                     /lab_host="Xl1-Blue"
                     /clone_lib="ApMS"
                     /note="Vector: pBS-SK minus; Site_1: EcoRI; Site_2: XhoI;
                     Sample name: ID0AAA ; Plant growth place: Department of
                     Ecology & Evolutionary Biology, Princeton University ;
                     Soil conditions: Soil ; Sowing date: 01/06/1999 ;
                     Harvesting date: 01/06/1999 ; Stress date: no stress ;
                     Description: Aphids inoculated on one-week old Vicia faba
                     under non-sterile conditions. All parthenogenetic stages
                     and both winged and wingless adults were collected for
                     library construction. ; experimental condition: long
                     photoperiod (16-hr light/8-hr dark at 18 c)"
ORIGIN

Query Match         34.2%;  Score 446.8; DB 8; Length 868;
Best Local Similarity   99.4%;
Matches    459;  Conservative   0;  Mismatches   2;  Indels   1;  Gaps   1;

Qy   1233 CATGTGTTCGTTGGCCAGGTTTTACGCAAAATCGTTGAATGAAGCACGCATACCGGCGCTGCAG 1292
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    475 CATGTGTTCGTTGGCCAGGTTTTACGCAAAATCGTTGAATGAAGCACGCATACCGGCGCGTGCAG 416

Qy   1293 TTTGAGAT 1300
          |||| |||
Db    415 TTTGTTAAT 408

Qy    839 AACGCCACGCCAGAGGTTGCCCGCAGCCGCCACCCCGTTATTAGATTTGATCAAGACAG 898
          ||||||||||||| ||||||||||||||||| |||||||||||||||||||||||||||
Db    868 AACGCCACGCCAGA-GTTGCCCGCAGCCGCCCACCCCGTTATTAGATTTGATCAAGACAG 810
```

RESULT 8

```
Qy    899 CGTTGACGCGCCCCATCCACCCGAAAAACAGGGGTATGGTGGTGTGACATTACCCACTTCAGTGC  958
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    809 CGTTGACGCGCCCCATCCACCCGAAAAACAGGGGTATGGTGGTGTGACATTACCCACTTCAGTGC  750

Qy    959 TGTTTATCGCCGGACACGATACTAATCTCGGGCGGCGGCGCGCACTGGAGCTCAACT  1018
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    749 TGTTTATCGCCGGACACGATACTAATCTCGGGCGGCGGCGCGCACTGGAGCTCAACT  690

Qy   1019 GGACGCGCTTCCCGGTCAGCCGGATAACACGCGCCCGCCCAGTGGGTGGAACGGCT  1078
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    689 GGACGCGTTCCCGGTCAGCCGGATAACACGCGCCCGCCCAGTGGGTGGAACGGCT  630

Qy   1079 GGCGTCGGGCTRAAGCGGATRAACAGCCAGTGGGATTCAGGTTTCGCTGGTGCTTCCGAGACTTTAC  1138
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    629 GGCGTCGGGCTRAAGCGGATRAACAGCCAGTGGGATTCAGGTTTCGCTGGTGCTTCCGAGACTTTAC  570

Qy   1139 AGCAGATGCGTGATAAAAACGGCCGCTGTCATTAAATACGGCCGCCCGTTATTAGA-TTGATCAAGACAGCGTT  1198
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db    569 AGCAGATGCGTGATAAAAACGGCCGCTGTCATTAAATACGGCCGCCCG............CAGCGTT  510

Qy   1199 CCCTGCGAGGATGTGAAGAGAGGGAAAATGCGCAGGGCACGGTATGGTGTGACAATTACCCACTTCAGTGCTGTT  1258
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    509 CCCTGCGAGGATGTGAAGAGAGGGAAAATGCGCAGGGCACGGTATGGTGTGACAATTACCCACTTCAGTGCTGTT  450

Qy   1259 AAAATCGTGAATGAAGCAGCACGCAATACCCGGCGTGCCAGTTTGAGAT  1300
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    449 AAAATCGTGAATGAAGCAGCACGCAATACCCGGCGTGCCAGTTTGTAAT  408
```

LOCUS   CN757565/c
DEFINITION ID0AA1DF09RM1 ApMS Acyrthosiphon pisum cDNA clone ID0AA1DF09 5', mRNA sequence.
                    CN757565   864 bp   mRNA   linear   EST 20-MAY-2004
ACCESSION   CN757565
VERSION   CN757565.1 GI:47531488
KEYWORDS   EST.
SOURCE   Acyrthosiphon pisum (pea aphid)
  ORGANISM Acyrthosiphon pisum
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes; Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
REFERENCE   1 (bases 1 to 864)
  AUTHORS   Hunter,W., Martinez-Torres,D., Rahbe,Y., Sabater-Munoz,B., Stern,D., Tagu,D. and Wincker,P.
  TITLE   An expressed sequence tags database for the pea aphid Acyrthosiphon pisum
  JOURNAL   Unpublished (2004)
COMMENT   Contact: D. Tagu
            INRA Rennes
            UMR BiO3P, BP 35327, F-35653 Le Rheu Cedex France
            Tel: +33.2.23.48.51.65
            Fax: +33.2.23.48.51.50
            Risk of contamination by bacterial sequences from obligatory (Buchnera) or facultative endosymbionts. These sequences were obtained in the frame of the International Consortium of Aphid Genomics in collaboration with Genoscope
            PCR Primers
            FORWARD: CAGGAAACAGCTATGACC
            Plate: 1 row: F column: 9.
FEATURES   Location/Qualifiers
     source   1..864
                     /organism="Acyrthosiphon pisum"
                     /mol_type="mRNA"
                     /organism="Acyrthosiphon pisum"
                     /cultivar="developmentstage"
                     /db_xref="taxon:7029"
                     /clone="ID0AA1DF09"
                     /tissue_type="whole insect"
                     /dev_stage="nymphs and adults (parthenogenetic females)"
                     /lab_host="XL1-Blue"
                     /clone_lib="ApMS"
                     /note="Vector: pBS-SK minus; Site_1: EcoRI; Site_2: XhoI;

ORIGIN

Sample name: ID0AAA ; Plant growth place: Department of Ecology & Evolutionary Biology, Princeton University ; Soil conditions: Soil ; Sowing date: 01/06/1999 ; Harvesting date: 01/06/1999 ; Stress date: no stress ; Description: Aphids inoculated on one-week old Vicia faba under non-sterile conditions. All parthenogenetic stages and both winged and wingless adults were collected for library construction.; experimental condition: long photoperiod (16-hr light/8-hr dark at 18 c)"

RESULT 9

Query Match 33.9%; Score 442.8; DB 8; Length 864;
Best Local Similarity 99.3%; Pred. No. 7.7e-128;
Matches 455; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```
Qy    843 CACGCCAGAGGTTGCCCGCCAGTGGCCAGCGCGGCGGCGGCACCCCGTTATTAGATTTGATCAAGACAGCGTT  902
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    864 CACGCCAGAGGTTGCCCGCCAGTGGCCAGCGCGGCGGCGGCACCCCGTTATTAGA-TTGATCAAGACAGCGTT  806

Qy    903 GACGCCCCATCCACCCGAAAAACAGGCGGTATGGTGTGACATTACCCACTTCAGTGCTGTT  962
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    805 GACGCCCCATCCACCCGAAAAACAGGCGGTATGGTGTGACATTACCCACTTCAGTGCTGTT  746

Qy    963 TATCGCGCCGGACACGATACTAATCTCGGGCGAAAATCTCGGCGGGCACTGGAGCTCAACTGGGAC  1022
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    745 TATCGCGCCGGACACGATACTAATCTCGGGCGAAAATCTCGGCGGGCACTGGAGCTCAACTGGGAC  686

Qy   1023 GCTTCCCGGTCAGCCGGATAACACGCGCCCGCCCAGTGGGTGGAACGGCTGCGTCGGGCTRAAGCGGATRAACAGCGCCGCCCCAGTGGGATTCAGGTTTCGCTGGTGCTTCCAGACTTTACAGCA  1142
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    625 TCGGGCTRAAGCGGATRAACAGCGCCGCCCCAGTGGGATTCAGGTTTCGCTGGTGCTTCCAGACTTTACAGCA  566

Qy   1143 GATGCGTGATAAAAACGGCCGCTGTCATTAAATACGGCCGCCCGGAGAGGAGGTGAAACTGACCGT  1202
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    565 GATGCGTGATAAAAACGGCCGCTGTCATTAAATACGGCCGCCCGGAGAGGAGGTGAAACTGACCCT  506

Qy   1203 GGCAGGATGTGAAGAGAGGGAAAATGCGCAGGGCACGGTATGGTGTGTTCGTTGGGCAGGTTTTTACGCAAAT  1262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    505 GGCAGGATGTGAAGAGAGGGAAAATGCGCAGGGCACGGTATGGTGTGTTCGTTGGGCAGGTTTTTACGCAAAT  446

Qy   1263 CCTGAATGAAGCACGCAATACCCGGCGTGCCAGTTTGAGAT  1300
          ||||||||||||||||||||||||||||||||||||||||
Db    445 CCTGAATGAAGCACGCAATACCCGGCGTGCCAGTTTGTAAT  408
```

LOCUS   BE520240/c
           BE520240   393 bp   mRNA   linear   EST 19-MAR-2001
DEFINITION M11B12STM Arabidopsis developing seed Arabidopsis thaliana cDNA clone M11B12 5', mRNA sequence.
ACCESSION   BE520240
VERSION   BE520240.1 GI:9778242
KEYWORDS   EST.
SOURCE   Arabidopsis thaliana (thale cress)
  ORGANISM Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1 (bases 1 to 393)
  AUTHORS   White,J.A., Todd,J., Newman,T., Focks,N., Girke,T., Martinez de Ilarduya,O., Jaworski,J.G., Ohlrogge,J. and Benning,C.
  TITLE   A new set of Arabidopsis expressed sequence tags from developing seeds. The metabolic pathway from carbohydrates to seed oil
  JOURNAL   Plant Physiol. 124 (4), 1582-1594 (2000)
  PUBMED   11115876
COMMENT   Contact: Benning, C
            Dept. of Biochemistry & Molecular Biology
            Michigan State University
            224 Biochemistry, Michigan State University, East Lansing, MI 48824, USA

Tel: 517 355 1609
Fax: 517 353 9334
Email: benning@msu.edu
Michigan State University DNA Sequencing Facility
Biological Resource Center, The Ohio State University, 309 Botany &
Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX:
614292029371.
614292060603 TEL:

```
FEATURES             Location/Qualifiers
     source          1..393
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /ecotype="Columbia"
                     /db_xref="taxon:3702"
                     /clone="M1B12"
                     /tissue_type="seed"
                     /dev_stage="5-13 days after flowering"
                     /clone_lib="E.coli"
                     /note="Organ: Developing seed; Vector: pBluescript SK-;
                     Site_1: EcoRI; Site_2: XhoII"
ORIGIN
Query Match   28.5%;  Score 373.4;  DB 7;  Length 393;
Best Local Similarity   96.7%;  Pred. No. 5.4e-106;
Matches  380;  Conservative  0;  Mismatches  13;  Indels  0;  Gaps  0;

Qy  319 ATTATTGCTGATCTCGACGAGCCTACCCGTAAAACAGGCGAAGCCTTCGCCGCCGCCGTG 378
        ||||||||||||||||||||||||| |||||||||||||||||| |||||||||||  ||
Db  393 ATTATTGCTGATCTCGACGAGCCTACCCGTAAAACAGGCGAAGCTTTCGCCGCCGGGNTG 334

Qy  379 GCACCTGACTGTGCAATAACCCTTACATACCCAGGCAGCAGATACGTTCCAGTCCCAGTTA 438
        |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db  333 GCACCTGACTGTGCAATAACCGTTACATACCCAGGCAGCAGATACGTTCCAGTCCCAGTTA 334

Qy  439 TTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCGGAACGTGACNCGGGATC 498
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  273 TTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAAACGCGGAACGTGACNCGGGATC 214

Qy  499 CTCGAGAGGGCAGGAGGGTCAATTGCTGACTTTACCCGGCATTATCAAACGGGGTTTCGC 558
        ||||||||||||||||||||||||||||||||||||||||         ||||||||||
Db  213 CTCGAGAGGGCAGGAGGGTCAATTGCTGACTTTACCCGGCAAACGGGGTTTCGC 154

Qy  559 GAACTGGAACGGGGTGCTTAATTTTCCCGGCAAATCAAACTTGTGGCTTAAATTTTCCGCA 618
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  153 GAACTGGAACGGGGTCAATTGCTTAAATTTTCCGGCAAATCAAACTTGTGCCTTAAACGTGAGAAACAG 94

Qy  619 GACGGAAAGCTGTTCATTAACGCAGGCCATTACCATGGGAACTCAAGGTGAGCGCCGACTGT 678
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   93 GACGGAAAGCTGTTCATTAACGCAGGCCATTACCATCGGAACTCAAGGTGAGCGCCGACAAT 34

Qy  679 GTCTCATTAACCGGTGCGGTAAGCCTCGCATCA 711
        |||||||||| || ||||||||||||||||||
Db   33 GTCTCATTAAGCCTGCGGTAAGCCTCGCATCA 1
```

```
RESULT 10
LOCUS       AW036132
DEFINITION  AW036132 tomato seed, TAMU Lycopersicon esculentum cDNA clone
            cLEB1E23 similar to periplasmic phosphoanhydride phosphohydrolase
            precursor, putative, mRNA sequence.
ACCESSION   AW036132
VERSION     AW036132.1  GI:5894811
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM    Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 383)
AUTHORS     Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,
            Holt,I.E., Liang,F., Upton,J., Ronning,C.M., Craven,M.B.,
            Fujii,C.Y., Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C.,
```

383 bp   mRNA   linear   EST 18-MAY-2001

Martin,G.B., Tanksley,S.D. and Giovannoni,J.

```
TITLE       Generation of ESTs from tomato seed tissue
JOURNAL     Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
3 prime sequence.
```

```
FEATURES             Location/Qualifiers
     source          1..383
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEB1E23"
                     /tissue_type="seeds"
                     /dev_stage="quiescent seed"
                     /lab_host="XL1-Blue MRF'"
                     /clone_lib="tomato seed"
                     /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
                     XhoI; clBB - Tomato Seed EST Library. Directionally cloned
                     cDNAs inserted into pBluescript SK(-) at 5' end with
                     EcoRI and 3' end with XhoI site."
ORIGIN
Query Match   28.3%;  Score 370.2;  DB 7;  Length 383;
Best Local Similarity   97.9%;  Pred. No. 5.5e-105;
Matches  375;  Conservative  0;  Mismatches  8;  Indels  0;  Gaps  0;

Qy  348 TAAAACAGGCGGAAGCCTTCGCCGCCGCCGTGGCACCTGACTGTGCAATAACCGTACATAC 407
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TAAAACAGGCGGAAGCCTTCGCCGCCGCCGTGGCACCTGACTGTGCAATAACCGTACATAC 60

Qy  408 CCAGGCAGATACGTCCCAGTCCCGATCCGTTGATTTTATTTAATTCCTCTAAAAAACTGGGTTTGCCA 467
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CCAGGCAGATACGTCCCAGTCCCGATCCGTTTATTTTTAATTCCTCTAAAAAACTGGGCGTTTGCCA 120

Qy  468 ACTGGATAACGCGGAACGTGACGCGGATCCTCGAGAGGGCAGGAGGGTCAATTGCTGA 527
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 ACTGGATAACGCGGAACGTGACTGACGCGATCCCTCGACGAGGGCAGGAGGGTCAATTGCTGA 180

Qy  528 CTTTACCCGGGCATTATCAAACGGGGTTTCGGGAACTGGAACGTGGCTGCTTAAATTTTCCGCA 587
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CTTTACCCGGGCAAACGGGGTTTCGGGAACTGGAACGGGGCCGGTTTGCGGAACTGAGAAACTGA 240

Qy  588 ATCAAAACTTGTGCCTTAAACGTGAGAGAAACAGGAGGACGAAAAGCTGTTCATTAACGCAGGCATT 647
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ATCAAAACTTGTGCCTTAAACGTGAGAGAAACAGGACGAAAAGCTGTTCATTAACGCAGGCATT 300

Qy  648 ACCATGGGAACTCAAGGTGAGCGCCGACTGTGTCTCATTAACCGGTGCGGTAAGCCTCGC 707
        |||||||||||||||||||||||||||||||||||||| || ||||||||||||||||
Db  301 ACCATCGGAACTCAAGGTGAGCGCCGACAAATGTCTCATTAAGCCTGCGGTAAGCCTCGC 360

Qy  708 ATCAAATGCTGACCGGAGATATTTC 730
        |||||||||||||||||||||||||
Db  361 ATCAATGCTGACCGGAGATATTTC 383
```

```
RESULT 11
LOCUS       AW036134
DEFINITION  AW036134 tomato seed, TAMU Lycopersicon esculentum cDNA clone
            cLEB1E23 similar to periplasmic phosphoanhydride phosphohydrolase
            precursor, putative, mRNA sequence.
ACCESSION   AW036134
VERSION     AW036134.1  GI:5894813
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM    Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
```

AW036134/c   383 bp   mRNA   linear   EST 18-MAY-2001

```
REFERENCE   1  (bases 1 to 383)
AUTHORS     Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,
            Holt,I.E., Liang,F., Upton,J., Ronning,C.M., Craven,M.B.,
            Fujii,C.Y., Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C.,
            Martin,G.B., Tanksley,S.D. and Giovannoni,J.
TITLE       Generation of ESTs from tomato seed tissue
JOURNAL     Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html

FEATURES             Location/Qualifiers
     source          1..383
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEB123"
                     /tissue_type="seeds"
                     /dev_stage="quiescent seed"
                     /lab_host="XL1-Blue MRF'"
                     /clone_lib="tomato seed"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     XhoI; cLEB - Tomato Seed EST Library. Directionally cloned
                     cDNAs inserted into pBlueScript SK(-) at 5' end with
                     EcoRI and 3' end with XhoI site."
ORIGIN          5 prime sequence.

Query Match             28.3%;  Score 370.2;  DB 7;  Length 383;
Best Local Similarity   97.9%;
Matches  375;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy        348 TAAAACAGGCGGAAGCCTTCGGCCGCCGGCGGCCTGGCACGTGAGCAATAACCGTACATAC 407
              |||||||||||||||||||||||| ||||||||||||||||||||||||| ||||||||||
Db        383 TAAAACAGGCGGAAGCCTTCGGCCGGCCGGCGGCCTGGCACGTGAGCAATAACCGTACATAC 324

Qy        408 CCAGGCAGATACGTCCAGTCCGATCCGTTATTTAATTCCTCTAAAAAACTGGCGTTTGCCA 467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 CCAGGCAGATACGTCCAGTCCGATCCGTTATTTAATTCCTCTAAAAAACTGGCGTTTGCCA 264

Qy        468 ACTGGATAAACGCGAACGTGACTGACCGATCCTCGAGAGGGCCAGGAGGGTCAATTGCTGA 527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 ACTGGATAAACGCGAACGTGACTGACCGATCCTCGAGAGGGCCAGGAGGGTCAATTGCTGA 204

Qy        528 CTTTACCGGGGCAATTATCAAACGGCGTTTCGGCAACTGGAACTGGGTGGCTTAATTTCCGCA 587
              ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        203 CTTTACCGGGGCATCCGCAAAACGGCGTTTCGGCAACTGGAACTGGGTGGCTTAATTTCCGCA 144

Qy        588 ATCAAACTTGTGCCTTAAACGTGAGAAACAGGACGAGGATGTTCATTAACCGAGCCATT 647
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 ATCAAAACTTGTGCCTTAAACGTGAGAAACAGGACGAGGATGTTCATTAACCGAGCCATT 84

Qy        648 ACCAATCGGAACTCAAGGTGACGCGCCGACTGTGTCTCATTAACCGGTGCGGTAAGCCTCGC 707
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 ACCAATCGGAACTCAAGGTGACGCGCCGACTGTGTCTCATTAACCGGTGCGGTAAGCCTCGC 24

Qy        708 ATCAATGCTGACGGAGATATTTC 730
              |||||||||||||||||||||||
Db         23 ATCAATGCTGACGGAGATATTTC 1
```

```
BU713769
LOCUS       BU713769               531 bp    mRNA    linear   EST 23-OCT-2003
DEFINITION  SJAABUG001 Adult SjC 7/94 Schistosoma japonicum cDNA similar to
            sp|P07102|PPA_ECOLI PERIPLASMIC APPA PROTEIN PRECURSOR [INCLUDES:
            PHOSPHOANHYDRIDE PHOSPHOHYDROLASE (PH 2.5 ACID PHOSPHATASE) (AP);
            6-PHYTASE ], mRNA sequence.
ACCESSION   BU713769
VERSION     BU713769.1  GI:28321125
KEYWORDS    EST.
SOURCE      Schistosoma japonicum
  ORGANISM  Schistosoma japonicum
            Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
            Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE   1  (bases 1 to 531)
AUTHORS     Hu,W., Yan,Q., Shen,D.K., Liu,F., Zhu,Z.D., Song,H.D., Xu,X.R.,
            Wang,Z.J., Rong,Y.P., Zeng,L.C., Wu,J., Zhang,X., Wang,J.J.,
            Xu,X.N., Wang,S.Y., Fu,G., Zhang,X.L., Feng,Z., Chen,Z. and Han,Z.G.
            McManus,D.P., Xue,C.L., Feng,Z.,
TITLE       Evolutionary and biomedical implications of a Schistosoma japonicum
            complementary DNA resource
JOURNAL     Nat. Genet. 35 (2), 139-147 (2003)
PUBMED      12973349
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn.

FEATURES             Location/Qualifiers
     source          1..531
                     /organism="Schistosoma japonicum"
                     /mol_type="mRNA"
                     /strain="Chinese (Anhui) strain"
                     /db_xref="taxon:6182"
                     /sex="Male and female"
                     /tissue_type="Whole body"
                     /dev_stage="Adult worms"
                     /lab_host="Mouse and rabbit"
                     /clone_lib="Adult SjC 7/94"
                     /note="Vector: Lambda ZAP-II XR; Site_1: EcoR I; Site_2:
                     XhoI I; Several hundred adult Schistosoma japonicum
                     (Anhui, P.R. China, strain), of mixed sex, were perfused
                     from the mesenteries of experimentally infected mice and
                     rabbits at the Queensland Institute of Medical Research,
                     Brisbane, Australia (QIMR), and stored for several months
                     in liquid nitrogen. Subsequently, mRNA was isolated at the
                     QIMR from lysates of these worms by oligo dT
                     chromatography, using a kit from Pharmacia. The mRNA was
                     then shipped to Clontech, Palo Alto, CA, USA, who
                     constructed a cDNA library. First strand synthesis was
                     primed with an oligo-dT-XhoI-primer and synthesized using
                     M-MLV reverse transcriptase. Second strand synthesis was
                     accomplished with RNase H and T4 DNA polymerase. The
                     double stranded cDNA was ligated to EcoRI linkers,
                     digested with EcoRI and XhoI, and ligated into the
                     phagemid vector lambda ZAP II XR. After construction of
                     this directional library by Clontech, it was returned to
                     the QIMR. During analysis of the library at the QIMR, we
                     have found that a small percentage, 2% to 3%, of the
                     clones contain inserts that appear to be highly homologous
                     to sequences from salmonoid fishes, as determined by
                     homology comparisons using BLAST and by Southern
                     hybridization analysis to genomic DNA from salmon (Sigma
                     Chemical Co., St. Louis, MO) under stringent washing
                     conditions. The remainder of the clones appear to contain
                     S. japonicum sequences."
ORIGIN

Query Match             26.7%;  Score 348.6;  DB 3;  Length 531;
Best Local Similarity   88.0%;
Matches  425;  Conservative    0;  Mismatches   54;  Indels    4;  Gaps    4;

Qy        670 GCCGACTGTGTCTCATTAACCGGTGCGGTAAGCCTCGCCATCAATGCTGACGGAGATATTT 729
              ||||||  ||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCGACAATGTCTCATTAACCGGTGCCGGTAAGCCTCGCCATCAATGCTGACGGAGATCGCGG 60

Qy        730 CTCCTGCAACAAGCACAGGGAGAATGCCCGAGCCGGGAAGGACTACACCGATTCACAC 789
              ||||||||||| || | ||||||||||||||||||||||||||||||||||||||||
Db         61 CTCCTGCAACAAGCACAGGGAATGCCCGAGCCGGGTGGGGAAGGACTACACCGATTCACAC 120

Qy        790 CAGTGGAACACCTTGCTAAGTTTGCATAACGGGCAAATTGATTTGCTACAACGCACGCCA 849
```

RESULT 13

```
LOCUS       CN759004                  706 bp    mRNA    linear   EST 20-MAY-2004
DEFINITION  ID0AAA24BC04RM1 ApMS Acyrthosiphon pisum cDNA clone ID0AAA24BC04
            5', mRNA sequence.
ACCESSION   CN759004
VERSION     CN759004.1  GI:47532927
KEYWORDS    EST.
SOURCE      Acyrthosiphon pisum (pea aphid)
  ORGANISM  Acyrthosiphon pisum
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
            Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
REFERENCE   1  (bases 1 to 706)
  AUTHORS   Hunter,W., Martinez-Torres,D., Rahbe,Y., Sabater-Munoz,B.,
            Stern,D., Tagu,D. and Wincker,P.
  TITLE     An expressed sequence tags database for the pea aphid Acyrthosiphon
            pisum
  JOURNAL   Unpublished (2004)
COMMENT     Contact: D. Tagu
            INRA Rennes
            UMR BiO3P, BP 35327, F-35653 Le Rheu Cedex France
            Tel: +33.2.23.48.51.65
            Fax: +33.2.23.48.51.50
            Risk of contamination by bacterial sequences from obligatory
            (Buchnera) or facultative endosymbionts. These sequences were
            obtained in the frame of the International Consortium of Aphid
            Genomics in collaboration with Genoscope
            PCR Primers
            FORWARD: CAGGAAAACAGCTATGACC
            Plate: 24  row: C  column: 4.
FEATURES             Location/Qualifiers
     source          1..706
                     /organism="Acyrthosiphon pisum"
                     /mol_type="mRNA"
                     /cultiva="developmentstage"
                     /db_xref="taxon:7029"
                     /clone="ID0AAA24BC04"
                     /tissue_type="whole insect"
                     /dev_stage="nymphs and adults (parthenogenetic females)"
                     /lab_host="XL1-Blue"
                     /clone_lib="ApMS"
                     /notes="Vector: pBS-SK minus; Site 1: EcoRI; Site_2: XhoI;
                     Sample name: ID0AAA ; Plant growth place: Department of
                     Ecology & Evolutionary Biology, Princeton University ;
                     Soil conditions: Soil ; Sowing date: 01/06/1999 ;
                     Harvesting date: 01/06/1999 ; Stress date: no stress ;
                     Description: Aphids inoculated on one-week old Vicia faba
                     under non-sterile conditions. All parthenogenetic stages
                     and both winged and wingless adults were collected for
                     library construction. ; experimental condition: long
                     photoperiod (16-hr light/8-hr dark at 18 c)"
ORIGIN

Query Match     22.6%;  Score 295.8;  DB 8;  Length 706;
Best Local Similarity   99.3%;  Pred. No. 2.1e-81;
Matches  297;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy   1002 CGACTGGAGCTCAACTGGACGCTTCCCGGTCAGCCGGATAACACGCCGGATAAC 1061
          |||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db    706 CGCACTGGAGCTCAACTGGACGCTTTCCCGGTCAGCCGGATAACACGCCGGATAAC 647

Qy   1062 ACTGGTGTTTGAACGCTGCGGTCGGCTAAGCGGATAACACGCCAGTGGATTCAGGTTTCGGCT 1121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    646 ACTGGTGTTTGAACGCTGCGGTCGGCTAAGCGGATAACACGCCAGTGGATTCAGGTTTCGGCT 587

Qy   1122 GGTCTTCCAGACATTTACAGCAGATGCGGTGGTGGATGCGTGATAAAACGCCC 1181
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    586 GGTCTTCCAGACATTTACAGCAGATGCGGTGG 527

Qy   1182 CGGAGGAGGTGAAACTGACCCTGGCAGGATGGATGTGAAGACGGAAATGCGGCAGGG 1241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    526 CGGAGGAGGTGAAACTGACCCTGGCAGGATGGATGTGAAGACGGAAATGCGGCAGGG 467

Qy   1242 GTTCGGCAGGTTTTACGCAAATCGTGAAATGAAAGCAACGCATACCGGCG 1300
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    466 GTTCGGCAGGTTTTACGCAAATCGTGAAATGAAAGCAACGCATACCGGCG 408
```

RESULT 14

```
LOCUS       AA545747                  354 bp    mRNA    linear   EST 12-MAY-1999
DEFINITION  HBMSP1B4-REV Human Bone Marrow Stromal Fibroblast Homo sapiens cDNA
            clone HBMSP1B4 5', mRNA sequence.
ACCESSION   AA545747
VERSION     AA545747.1  GI:2307026
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 354)
  AUTHORS   Jia,L., Robey,P., Young,M., Schuler,G., Powell,J., Yang,L.,
            Lennon,G., Hillier,L., Allen,M., Bowles,L., Geisel,S., Kucaba,T.,
            Marra,M., Martin,J., Steptoe,M., Tan,F., Theising,B., Bowers,Y.,
            Wylie,T., Waterston,R., Wilson,R. and Franconamo,C.
  TITLE     Washu-MGB/NHGRI EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Libin Jia
            National Human Genome Research Institute
            Medical Genetics Branch
            10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
            Tel: 301-402-4877
            Fax: 301-496-7157
            Email: libin@helix.nih.gov
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1..354
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HBMSP1B4"
                     /sex="Male and Female"
                     /tissue_type="bone marrow"
                     /cell_type="stromal fibroblast"
                     /dev_stage="mixed"
                     /lab_host="XL1-Blue"
```

```
Db   121 CAGTGGGACACCTTGCTAAGTTTGCATAAGCGGCAAATTTATTTGTACTACAACGGCACGCCCA 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   850 GAGGTTGCCGCAGCGCGGCCGCCACCCCGTTTATTTAGATTTGATCAAGACAGCGTTGACGGCCC 909
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GAGGTTGCCGCAGCCGCGCCACCCCGGTATTAGATTTGATCAAGACACCGTTGACGCCC 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   910 CATCCACCGGCAAAACAGGCGTATGGCGTGTGACATTACCCACTTCAGTG-CTGTTTATCGC 968
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CGTTTATCATTATGATGTTGCATCTGTGAGAGACAATTCCCTCTTGAGTGACGAAAAAGG 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   969 CGGACACGATACTAATCTCGGCAAAATCTCGGCGGCCACTGGAGCTCAACTGGACGCTTCC 1028
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CGGCCACCACGAGCCTGAGCTGATAAAATCTGGGCGGCCACTGGAGCTCAACTGGACGCTGTC 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1029 CGGTCAGCCGGATAACACGCCGCCGGATGGTGGTTGGAACTGG-TGTTTGAACGCT-GGCGTCGG 1086
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CAGGCGAGCCCGATAACACGCCCCGCGAGGACCTGAATGTTTGAACCCTGGCCGGTCGG 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1087 CTAAGCCGATAACAGCCACTGGATTCAGGTTTCGGCTGGTTCTTCCAGACTTTACAGCAGATG 1146
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CTAAGCCGATAACAGCCACTGGATTCAGGTTTCGGTGTCCTGGGTCTTCCAGAC-TTACAGCAGATG 479
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1147 CGT 1149
         |||
Db   480 CGT 482
```

```
ORIGIN
        /clone_lib="Human Bone Marrow Stromal Fibroblast"
        /note="Vector: pBluescript; Site_1: EcoRI; Site_2: XhoI"

Query Match          22.6%;    Score 295.4;   DB 1;  Length 354;
Best Local Similarity  91.4%;    Pred. No. 2.4e-81;
Matches 322; Conservative   0; Mismatches   28; Indels   1; Gaps   1;

Qy    499 CTCGAGGGGCAGGAGGGTCAATTGCTACTTACCGGGCAGCAATATCAAACGCGGTTTCGC 558
          ||||||||||||||||||||||||||||||||| ||||||||| |||||||||||||||||
Db    350 CTTCAGCAGCAGGCAGCAGAGGTCAAATTGCTGATTATTACCGGGCAGCAAACGCGTTTCGC 291

Qy    559 GAACTGGAACGGGGTGCTTAATTTTCCGCAATCAAACTTGTGCCTTAAAACGTGAGAAACAG 618
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    290 GAACTGGAACGGGGTG-TTAATTTTCCGCAATCAAACTTGTGCCTTAAAACGTGAGAAACAG 232

Qy    619 GACGAAAAGCTGTTCATTAACCGAGGGCAATTACCCATCGGAACTTCAAGGTGAGGGCCGAC 678
          |||||||||||||||||||||||||||||||||||| |||||||||| |||||||||||||
Db    231 GACGAAAAGCTGTTCATTAACCCAGGGCAATTACCATCGGAACTCAAGGTGAGGGCCCGAC 172

Qy    679 GTCTCAATTAACCGGTGGGGTAAGCCTGGCATCAAATGCTGACGGAGATATTTCTCCTGCA 738
          ||||| |||| |||||||||||||||||||||||| ||||||||||||||||||||||||
Db    171 GTNTCAATTAACCGGTCCGGTAGCCTTCGGCATCAAATGCTGACGGAGATATTNTCCTGCAA 112

Qy    739 CAAGCACAGGGAAATGCCGGACCCGGGTGGGGAAGCATCACCGATTCACCAGTGGAAC 798
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    111 CAAGCACAGGGAAATGCCGGACGCGGGGTGGGGAAGGATCACCGATTCACCCAGTGGAAC 52

Qy    799 ACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAACGCACGCCA 849
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    51  ACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAACGCACGCCA 1
```

```
RESULT 15
BJ619443
LOCUS       BJ619443          707 bp    mRNA       linear       EST 01-OCT-2003
DEFINITION  BJ619443 NIBB Mochii normalized Xenopus early gastrula library
            Xenopus laevis cDNA clone XL189d24 5', mRNA sequence.
ACCESSION   BJ619443
VERSION     BJ619443.1  GI:37258203
KEYWORDS    EST.
SOURCE      Xenopus laevis (African clawed frog)
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Xenopus.
REFERENCE   1  (bases 1 to 707)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and
            Kohara,Y.
  TITLE     Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
  COMMENT   Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp
            The information of this clone is available through the following
            URL.
            http://xenopus.nibb.ac.jp.
FEATURES             Location/Qualifiers
     source          1..707
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8355"
                     /clone="XL189d24"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 10.5"
                     /clone_lib="NIBB Mochii normalized Xenopus early gastrula
                     library"
ORIGIN
```

```
Query Match          19.4%;    Score 254.4;   DB 2;  Length 707;
Best Local Similarity  99.6%;    Pred. No. 2.6e-68;
Matches 255; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy    750 AATGCCCGGAGCCGGGGGTGGGGAAGGAATCACCGATTCACACCAGTGGAACACCTTGCTAAG 809
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1   AATGCCCGGAGCCGGGGGTGGGGAAGGAATCACCGATTCACACCAGTGGAACACCTTGCTAAG 60

Qy    810 TTTGCATAACGCGCAATTTGATTGTTGCTACAACGCACGCCAGAGGTTGCCCGCCGCGCC 869
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  TTTGCATAACGCGCAATTTATTTGCTACAACGCACGCCAGAGGTTGCCCGCCGCGCC 120

Qy    870 CACCCCGTTATTAGATTTGATCAAGACAGCGGTTGACGCCCCATCCACCGCAAAAACAGGC 929
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CACCCCGTTATTAGATTTGATCAAGACAGCGGTTGACGCCCCATCCACCGCAAAAACAGGC 180

Qy    930 GTATGGTGTGACATTACCCACTTCAGTGTCTGTTTTATCGGCCGGGACACGGATACTAAT 989
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTATGGTGTGACATTACCCACTTCAGTGTCTGTTTTATCGGCCGGGACACGGATACTAAT 240

Qy    990 AAATCTCGGCGCGCGCA 1005
          |||||||||||||||||
Db    241 AAATCTCGGCGCGCGCA 256

Search completed: June 14, 2006, 14:21:52
Job time : 6872.62 secs
```

OM nucleic - nucleic search, using sw model

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

Run on: June 14, 2006, 10:29:05 ; Search time 1596.84 Seconds
                 (without alignments)
                 10064.996 Million cell updates/sec

Title: US-10-601-319-9
Perfect score: 1308
Sequence: 1 atgaaagcgatcttaatccc...........gcagttcgagatctcatcta 1308

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0
Searched: 18892170 seqs, 614381763B residues

Total number of hits satisfying chosen parameters: 37784340

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Database :   Published Applications NA Main:*
                 1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                 2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                 3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
                 4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
                 5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
                 6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09D_PUBCOMB.seq:*
                 7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09E_PUBCOMB.seq:*
                 8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09F_PUBCOMB.seq:*
                 9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
                 10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
                 11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
                 12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
                 13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
                 14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq:*
                 15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq:*
                 16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq:*

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1308	100.0	1308	6	US-10-156-660-1	Sequence 1, Appli
2	1308	100.0	1308	8	US-10-601-319-9	Sequence 9, Appli
3	1308	100.0	10	US-10-933-115-9	Sequence 9, Appli
4	1296	99.1	1296	14	US-11-056-354-1	Sequence 1, Appli
5	1268.6	97.0	1296	8	US-10-601-319-6	Sequence 6, Appli
6	1267	96.9	1901	8	US-10-601-319-5	Sequence 5, Appli
7	1267	96.9	1901	10	US-10-601-319-5	Sequence 5, Appli
8	1265	96.7	1901	3	US-10-933-115-5	Sequence 1, Appli
9	1265	96.7	1323	3	US-09-777-566A-1	Sequence 1, Appli
10	1265	96.7	1323	5	US-09-866-379-1	Sequence 1, Appli
11	1265	96.7	1323	7	US-10-034-985-1	Sequence 1, Appli
12	1265	96.7	1323	8	US-10-430-356-1	Sequence 1, Appli
13	1265	96.7	1323	10	US-10-601-319-1	Sequence 1, Appli
14	1264	96.6	1299	8	US-10-282-122A-7167	Sequence 7167, Ap
15	1263.8	96.6	1901	8	US-09-866-379-7	Sequence 6, Appli
16	1263.8	96.6	1901	3	US-09-866-379-7	Sequence 7, Appli
17	1263.8	96.6	1901	6	US-10-156-660-3	Sequence 3, Appli
18	1263.8	96.6	1901	8	US-10-601-319-7	Sequence 7, Appli
19	1263.8	96.6	1901	10	US-10-933-115-6	Sequence 6, Appli
20	1263.8	96.6	1901	10	US-10-933-115-7	Sequence 7, Appli
21	1263.8	96.6	1901	14	US-11-056-354-3	Sequence 3, Appli
22	1262.2	96.5	1901	3	US-09-866-379-5	Sequence 5, Appli
23	1262.2	96.5	1901	8	US-10-601-319-6	Sequence 6, Appli
24	1233.6	94.3	1486	13	US-11-018-709-2	Sequence 2, Appli
25	1228	93.9	1489	6	US-10-266-041-9	Sequence 9, Appli
26	1228	93.9	1489	7	US-10-284-962-1	Sequence 1, Appli
27	1222.4	93.5	1486	7	US-10-284-962-4	Sequence 4, Appli
28	1222.4	93.5	1486	13	US-11-018-709-4	Sequence 4, Appli
29	758.2	58.0	11357	13	US-11-074-522-14	Sequence 14, Appl
30	691.4	52.9	1281	7	US-10-334-671-4	Sequence 4, Appli
31	691.4	52.9	1281	10	US-10-450-763-25689	Sequence 25689, A
32	691	52.8	1737	10	US-10-450-763-25689	Sequence 25689, A
33	404.8	30.9	466	8	US-10-282-122A-31181	Sequence 3181, Ap
34	377.2	28.8	1264	10	US-10-450-763-24247	Sequence 24247, A
35	377.2	28.8	1264	10	US-10-450-763-30120	Sequence 30120, A
36	377	28.8	4050	10	US-10-450-763-386	Sequence 386, App
37	230.2	17.6	1326	8	US-10-282-122A-41608	Sequence 41608, A
38	226.2	17.3	1325	8	US-10-021-723A-3	Sequence 3, Appli
39	223	17.0	1325	6	US-10-021-723A-1	Sequence 1, Appli
40	212.4	16.2	1326	7	US-10-021-723A-11	Sequence 11, Appl
41	191	14.6	356	10	US-10-450-763-25687	Sequence 25687, A
42	164.2	12.6	254	7	US-10-317-444-301	Sequence 301, App
43	164.2	12.6	254	7	US-10-317-444-302	Sequence 302, App
44	164.2	12.6	254	7	US-10-317-444-303	Sequence 303, App
45	164.2	12.6	254	7	US-10-317-444-304	Sequence 304, App

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1
US-10-156-660-1

Sequence 1. Application US/10156660
Publication No. US20030103958A1
GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson R.
; APPLICANT: Garrett, James B.
; APPLICANT: O'Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM
; TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 09010-029007
; CURRENT APPLICATION NUMBER: US/10/156,660
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
;
; SEQ ID NO 1
; LENGTH: 1308
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified phytase enzyme
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1308)

```
US-10-156-660-1

Query Match             100.0%;  Score 1308;  DB 6;  Length 1308;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1308;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAAACTGCA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAAACTGCA 60

Qy   61 TTCGCTTCAGAGTGAGCGGGAGCTGAAAGCTGAAAGTGTGGTGATTGGTCAGTTCATGGT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTTCAGAGTGAGCGGGAGCTGAAAGCTGAAAGTGTGGTGATTGGTCAGTTCATGGT 120

Qy  121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGCAGTGTCACCCCGAGCCATGGCCA 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGCAGTGTCACCCCGAGCCATGGCCA 180

Qy  181 ACCTGCGCCGGTAAAACTGGGGTGAGCTGACACCCGGCGGCGGGGTGGGTGAGCTTATCTC 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ACCTGGCCGGTAAAACTGGGGTGAGCTGACACCCGGCGGCGGGGTGGGTGAGCTTATCTC 240

Qy  241 GGACATTACTGGGCGTCAGCGTTCGGTAGCCGACGGATTGCTGCCCTAAAATGGGCTGCCCG 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GGACATTACTGGGCGTCAGCGTTCGGTAGCCGACGGATTGCTGCCCTAAAATGGGCTGCCCG 300

Qy  301 CAGTCTGGTCAGGTCGGCGATTATTGCTGATGTGACGACGTACCGTAAAAACAGGCGAA 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGTCTGGTCAGGTCGGCGATTATTGCTGATGTGACGACGTACCGTAAAAACAGGCGAA 360

Qy  361 GCCTTCGCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GCCTTCGCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420

Qy  421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAAACTGGCGTGGGATTAACGGCG 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAAACTGGCGTGGGATTAACGGCG 480

Qy  481 AACGTGACTGACGCGATCCTCGAGACGGGCACGAGGGGTCAATGCTGACTTTACCGGGCAT 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 AACGTGACTGACGCGATCCTCGAGACGGGCACGAGGGGTCAATGCTGACTTTACCGGGCAT 540

Qy  541 TATCAAAACGGCGTTTCGCGAACTGGAACGCGGGTGCTTAATTTTCCCGCAATCAAAACTTGTGC 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 TATCAAAACGGCGTTTCGCGAACTGGAACGCGGGTGCTTAATTTTCCCGCAATCAAAACTTGTGC 600

Qy  601 CTTAAAACGTGAGAAAACAGGACGAAAAGCTGTTCATTAACGCGAATTACCATCGGAACTC 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CTTAAAACGTGAGAAAACAGGACGAAAAGCTGTTCATTAACGCGAATTACCATCGGAACTC 660

Qy  661 AAGGTGAGCGCGACTGTGTCTCTTATTAACCGGTGAGCCTCGGCCATCAAATGCTGACG 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 AAGGTGAGCGCGACTGTGTCTCTTATTAACCGGTGAGCCTCGGCCATCAAATGCTGACG 720

Qy  721 GAGATATTTCTCCTCGCAACAAGCACAGGGAATGCCGGACGGAAGGAATCACC 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GAGATATTTCTCCTCGCAACAAGCACAGGGAATGCCGGACGGAAGGAATCACC 780

Qy  781 GATTCACACAGTGAACACCTTGCTAAGTTTGCATAACGGGCAAATTTGATTTGCTACAA 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GATTCACACAGTGAACACCTTGCTAAGTTTGCATAACGGGCAAATTTGATTTGCTACAA 840

Qy  841 CGCACGCCCAGAGGTTGCCCGCGCCAGCCGCGCCACCCGTTATTAGATTGATCAAGACAGCG 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 CGCACGCCCAGAGGTTGCCCGCGCCAGCCGCGCCACCCGTTATTAGATTGATCAAGACAGCG 900

Qy  901 TTGACGCCCCATCCACCGCAAAAAACAGGGCGTATGGTGTGACATTACCCACTTCAGTGCTG 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TTGACGCCCCATCCACCGCAAAAAACAGGGCGTATGGTGTGACATTACCCACTTCAGTGCTG 960

Qy  961 TTTACCGCCCGGACACGACGATACTTAATCTGGCGAAAATCTCGGGCGCGGCCGCACTCTGGAGGCTCAACTGG 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 TTTACCGCCCGGACACGACGATACTTAATCTGGCGAAAATCTCGGGCGCGGCCGCACTCTGGAGGCTCAACTGG 1020
```

```
Qy 1021 ACGCTTCCCGGTCAGCCGGGATAAACACGCCCGCGCCCAGGGTGGGTGAAACCGCTGG 1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 ACGCTTCCCGGTCAGCCGGGATAAACACGCCCGCGCCCAGGGTGGGTGAAACCGCTGG 1080

Qy 1081 CGTCGGCTAAGCGGATAACAGCCCAGTGGAATTCAGGTTTCGCGTGGTGTTTGAAACCGCTGG 1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 CGTCGGCTAAGCGGATAACAGCCCAGTGGAATTCAGGTTTCGCGTTCTTCCAGACTTTACAG 1140

Qy 1141 CAGATGCGTGATAAAAACGACCCCGGTCATTAAATCACGCGCCCGGGGAGAGGAGAGAAATGACC 1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 CAGATGCGTGATAAAAACGACCCCGGTCATTAAATCACGCGCCCGGGGAGAGGAGAGAAATGACC 1200

Qy 1201 CTGGCAGGATGTGAACGAGAGGCAGAGCAGGGCCATGTGTCGGTTGGTTCGTTTTTACGGCAA 1260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 CTGGCAGGATGTGAAGAGAGCGAAATGCCGCAGGGCATGTGTCGGTTGGTTCGTTTTTACGGCAA 1260

Qy 1261 ATCGTGAATGAAGCACGCGCCATACCCGGCGTGCAGTTTGAGATCTCATCTA 1308
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 ATCGTGAATGAAGCACGCGCCATACCCGGCGTGCAGTTTGAGATCTCATCTA 1308
```

```
RESULT 2
US-10-601-319-9

; Sequence 9, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
;  APPLICANT: Short, Jay M.
;  APPLICANT: Kretz, Keith A.
;  APPLICANT: Gray, Kevin A.
;  APPLICANT: Barton, Nelson Robert
;  APPLICANT: Garrett, James B.
;  APPLICANT: O' Donoghue, Eileen
;  APPLICANT: Mathur, Eric J.
;  TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
;  TITLE OF INVENTION:  AND USING THEM
;  FILE REFERENCE: 09010-029011
;  CURRENT APPLICATION NUMBER: US/10/601,319
;  CURRENT FILING DATE:  2003-06-20
;  PRIOR APPLICATION NUMBER: US 09/866,379
;  PRIOR FILING DATE: 2001-05-24
;  PRIOR APPLICATION NUMBER: US 09/580,515
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: US 09/318,528
;  PRIOR FILING DATE: 1999-05-25
;  PRIOR APPLICATION NUMBER: US 09/291,931
;  PRIOR FILING DATE: 1999-04-13
;  PRIOR APPLICATION NUMBER: US 09/259,214
;  PRIOR FILING DATE: 1999-03-01
;  PRIOR APPLICATION NUMBER: US 08/910,798
;  PRIOR FILING DATE: 1997-08-13
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: FastSEQ for Windows Version 4.0
;
; SEQ ID NO 9
;  LENGTH: 1308
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: modified phytase enzyme
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (1)...(1296)
US-10-601-319-9

Query Match             100.0%;  Score 1308;  DB 8;  Length 1308;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1308;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAAACTGCA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAAACTGCA 60

Qy   61 TTCGCTTCAGAGTGAGCGGGAGCTGAAAGCTGAAAGTGTGGTGATTGGTCAGTTCATGGT 120
```

RESULT 3
US-10-933-115-9
; Sequence 9, Application US/10933115
; Publication No. US20050281792A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 564462001822
; CURRENT APPLICATION NUMBER: US/10/933,115
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn for Windows Version 4.0
;
; SEQ ID NO 9
; LENGTH: 1308
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified phytase enzyme
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1296)
US-10-933-115-9

Query Match   100.0%; Score 1308; DB 10; Length 1308;
Best Local Similarity   100.0%;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGAAAGCGATCTTAATCCCATTTTTATATCTTCTTCGTTAAACCCGCCAATCTGCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAGCGATCTTAATCCCATTTTTATATCTTCTTCGTTAAACCCGCCAATCTGCA 60

Qy   61 TTCGCTCAGAGTGAGCCCGGAGCTGAAAGTTGGAAAAGTGTCGTGTCGTCATGGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCCGGAGCTGAAAGTTGGAAAAGTGTCGTGTCGTCATGGT 120

Qy   121 GTGCGTGCTCCAACCAAGGCCACGGACGATCGATCCGAGATGTCACCCGCAATCGCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTGCGTGCTCCAACCAAGGCCACGGACGATCGATCCGAGATGTCACCCGCAATCGCA 180

Qy   181 ACCTGGGCCGGGTAAAAACTGGGTGGGTGAGCTGAACCCCGGGTGGCTAATCTCATCTA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ACCTGGGCCGGGTAAAAACTGGGTGGGTGAGCTGAACCCCGGGTGGCTAATCTCATCTA 240

RESULT 4
US-11-056-354-1

Query Match   99.1%; Score 1296; DB 14; Length 1296;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCGTTAACCCCGCAAATCTGCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCGTTAACCCCGCAAATCTGCA 60

Qy     61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTTGAAAAGTGTGGTGATTGTCAGTGTCATGGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTTGAAAAGTGTGGTGATTGTCAGTGTCATGGT 120

Qy    121 GTGGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGAGTCACCCCAGACGGCCATGGCCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGAGTCACCCCAGACGGCCATGGCCA 180

Qy    181 ACCTGGCCCGGTAAAACTGGGGTGAGCTGACACCGCGCGCGTGGTAATCGCCTATCTC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACCTGGCCCGGTAAAACTGGGGTGAGCTGACACCGCGCGCGTGGTAATCGCCTATCTC 240

Qy    241 GGACAATTACTGCGTCGACGGTCGCGGATTATTGCTGGATGTCGACACGGAGTTGCTGCCCG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGACAATTACTGCGTCGACGGTCGCGGATTATTGCTGGATGTCGACACGGAGTTGCTGCCCG 300

Qy    301 CAGTCTGGTCAGGTCGCGGATTATTGCTGGATGTCGACACGGAGTTGCTGCCCG 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGTCTGGTCAGGTCGCGGATTATTGCTGGATGTCGACACGGAGTTGCTGCCGAA 360

Qy    361 GCCTTCGCCGCCGCGGGGCTGGCCACCTGACTGTGCAATAACCGTACTACCCAGGCAGDATACG 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCTTCGCCGCCGCGGGGCTGGCCACCTGACTGTGCAATAACCGTACTACCCAGGCAGATACG 420
```

```
Qy    241 GGACAATTACTGCGGTCAGGTCTCGGTAGCCGGACGGATTGCTGCCCTAAAATGTGGGCTGCCCG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGACAATTACTGCGGTCAGGTCTCGGTAGCCGGACGGATTGCTGCCCTAAAATGTGGGCTGCCCG 300

Qy    301 CAGTCTGGTCAGGGTCAGGGTCGCGGATTATTGCTGTGATGTCGACGAGCGGAGCCGAA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GCCTTCGCCGCCGCGGGGCTGGCCACCTGACCTGACTGTGCAATAACCGAGACAGATACG 360

Qy    361 GCCTTCGCCGCCGCGGGGCTGGCCACCTGACTGTGCAATAACCGTGCAATAACCAGGCAGDATACG 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGGCGTTTGCCAACTCGGATAACGCG 480

Qy    421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGGTTTGCCAACTCGGATAACGCG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AACGTGACTGACGCGGATCCTCGAGAGGGCACAGGAGGGTCAATTGCTGACTTTACCGGGCAT 540

Qy    481 AACGTGACTGACGCGGATCCTCGAGAGGGCACAGGAGGGTCAATTGCTGACTTTACCGGGCAT 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TATCAAACGGGCGTTTGGCGAACTGGAACGGGTGCTTAAATTTTCCGGCAATCAAACTTGTGC 600

Qy    541 TATCAAACGGGCGTTTGGCGAACTGGAACGGGTGCTTAAATTTTCCGGCAATCAAACTTGTGC 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTTAAACGTGAGAAACAGGAGCGAAACTGTTCATTCATTCCGGCAATCAATCGGAAGTCTGAC 660

Qy    601 CTTAAACGTGAGAAACAGGAGCGAAACTGTTCATTCATTCCGGCAATCAATCGGAAACTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AAGGTGAGCGCCGACTGTGTCTCATTAACCGGTGCCGGTAAGCCTCGCCATCAATGCTGACG 720

Qy    661 AAGGTGAGCGCCGACTGTGTCTCATTAACCGGTGCCGGTAAGCCTCGCCATCAATGCTGACG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GAGATATTTCTCTGCAACAAGCACAGGGAAATGCCGGACGCCGGGTGTGGGGAAGGATCACC 780

Qy    721 GAGATATTTCTCTGCAACAAGCACAGGGAAATGCCGGACGCCGGGTGGGGAAGGATCACC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCCCAATTGATTGCTACAA 840

Qy    781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCCCAATTGATTGCTACAA 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CGCACCCCAGAGGTTGCCCCGCCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 900

Qy    841 CGCACCCCAGAGGTTGCCCCGCCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCTG 960

Qy    901 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTTATCGCCGGACACGGATACTAATCTGGCAAAATCTCGGCGGCCACTGGAGCTCAACTGG 1020

Qy    961 TTTATCGCCGGACACGGATACTAATCTGGCAAAATCTCGGCGGCCACTGGAGCTCAACTGG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGCTTCCCGGTCAGCCGGATAACAGCCCGCGGGATAACGCTGTTGGAACCGCTGG 1080

Qy   1021 ACGCTTCCCGGTCAGCCGGATAACAGCCCGCGGGATAACGCTGTTGGAACCGCTGG 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CGTCCGGCTAAGCGATAACAGCCAGTGGAATTCAGGTTTGCTTCTTCCAGACTTTACAG 1140

Qy   1081 CGTCCGGCTAAGCGATAACAGCCAGTGGAATTCAGGTTTGCTTCTTCCAGACTTTACAG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAGATGCGTGATAAAAACGGCCGCTGTCAGTTCATTAAATACCGCCGGAGAGGTGAAAACTGACC 1200

Qy   1141 CAGATGCGTGATAAAAACGGCCGCTGTCAGTTCATTAAATACCGCCGGAGAGGTGAAAACTGACC 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGGCAGGATGTGAAGACGGGAAATGGGCAGGGGCCATGTGTTGTTCGGTGCAGGTTTTACGGCAA 1260

Qy   1201 CTGGGCAGGATGTGAAGACGGGAAATGGGCAGGGGCCATGTGTTGTTCGGTGCAGGTTTTACGGCAA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAGCACGCATACCGGGCGGTGCGCAGTTTGAGATCTCATCTA 1308

Qy   1261 ATCGTGAATGAAGCACGCATACCGGGCGGTGCGCAGTTTGAGATCTCATCTA 1308
```

; US-10-601-319-6
; Sequence 6, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING THEM
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 403
; OTHER INFORMATION: n = A,T,C or G
US-10-601-319-6

RESULT 5

Query Match 97.0%; Score 1268.6; DB 8; Length 1901;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

RESULT 6
US-10-601-319-5
; Sequence 5, Application US/10601319
; Publication No. US20040091968A1
GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING THEM
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931

; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 5
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 403
; OTHER INFORMATION: n = A,T,C or G
US-10-601-319-5

Query Match   96.9%;   Score 1267;   DB 8;   Length 1901;
Best Local Similarity   98.4%;   Pred. No. 0;   Mismatches 21;   Indels 0;   Gaps 0;
Matches 1279;   Conservative 0;   Mismatches 21;   Indels 0;   Gaps 0;

```
Qy      1  ATGAAAAGCGATCTTAATCCCCATTTTTATTCTCTTCTGATTCGTTAACCCCGCAAATCGCA   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    188  ATGAAAAGCGATCTTAATCCCCATTTTTATTCTCTTCTGATTCGTTAACCCCGCAAATCTGCA 247

Qy     61  TTCGCTCAGAGTGAGCCGGAGACTGAAGCTGAAAAAGTGTGGATTGTTCAGTCGTCCATGGT  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248  TTCGCTCAGAGTGAGCCGGAGACTGAAGCTGAAAAAGTGTGGATTGTTCAGTCGTCCATGGGT 307

Qy    121  GTGCGTTGCTCAACCAARGGCCACGCAACGTGATTCAGGAGTCACCCCAGACGCATGGGCCA  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308  GTGCGTTGCTCAACCAAGGCCACGCAACGTGATCAGGAGTCACCCCAGACGCATGGGCCA    367

Qy    181  ACCTGGCCGGTTAAAACTGGGTGAGCTGACACCGCGCGGTGGTGAGCTAATCGCTATCTC    240
           |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db    368  ACCTGGCCGGTTAAAACTGGGTGAGCTGACACCGCGNGGTGGTGAGCTAATCGCTATCTC    427

Qy    241  GGACATTACCAACGCCAGCGTCTGGTTAGCCGACGGATTGCTGGACATTGCTGCAATGTGGGT  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428  GGACATTACCAACGCCAGCGTCTGGTTAGCCGACGGATTGCTGGACATTGCTGCAATGTGGGT 487

Qy    301  CAGTCTGGTCAGGTCGGCGATTATTGCTGATGATGATCGAGAGCGGTACCCGTAAAAACAGGCGAA  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488  CAGTCTGGTCAGGTCGGCGATTATTGCTGATGATCGAGAGCGGTACCCGTAAAAACAGGCGAA  547

Qy    361  GCCTTGCCGCCGCGGGGTGGCACCTGACTGTGTCCAATAACCGTACATACCCAGGCAGATACG  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548  GCCTTGCCGCCGCGGGGTGGCACCTGACTGTGTCCAATAACCGTACATACCCAGGCAGATACG  607

Qy    421  TCCAGTCCCGATCCGGTTATTTAATCCTCTAAAAAACTGGCGTTTGCCAACTGGATAACGCG  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608  TCCAGTCCCGATCCGGTTATTTAATCCTCTAAAAAACTGGCGTTTGCCAACTGGATAACGCG  667

Qy    481  AACGTGACTGACGCGGATCCTCGAGAGGGCAGGAGGGTCAATGGCTTTACCGGGGCAT    540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    668  AACGTGACTGACGCGGATCCTCGAGAGGGCAGGAGGGTCAATGGCTTTACCGGGGCAT    727

Qy    541  TATCAAACGGGCTTTCGGACTGAAACGGGCTGCTTAATTTTCCGCAATCAAAACTTGTGC    600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    728  CGGCAAACGGGCGTTTCGCGACTGAAACGGGGTGCTTAATTTTCCGCAATCAAAACTTGTGC  787

Qy    601  CTTAAACGTGAGAAACAGGACGAAAAGCTGTTCATTAACGCAGGCAGCATTACCATCGGAACTC  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    788  CTTAAACGTGAGAAACAGGACGAAAAGCTGTTCATTAACGCAGGCAGCATTACCATCGGAACTC  847

Qy    661  AAGGTGAGCGCGGACCGACTGTGTCTCATTAACCGGTGCGGTAAGCCTCGGCATCAAATGCTGACG  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    848  AAGGTGAGCGCGGACCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGGCATCAAATGCTGACG  907

Qy    721  GAGATATTTCTCCTGCAACAAGCACAGGGAAATGCCGAGCCGGGGTGGGGAGGATCACC  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908  GAGATATTTCTCCTGCAACAAGCACAGGGAAATGCCGAGCCGGGGTGGGGAGGATCACC  967
```

```
Qy   781  GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCCCAATTTGAATTGATTTGGCTACA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   968  GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTGATTTGATTTATTTGCTACA  1027

Qy   841  CGCACGCGCCAGAGGTTGCCCGCGACCAGCGCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1028  CGCACGCGCAGAGGTTGCCCGCGACCAGCGCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG  1087

Qy   901  TTGACGACGCCCATCCACCCGCAAAAAACAGGCGTATGGGTGTGACATTACCCACTTCAGTGCTG  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1088  TTGACGACGCCCATCCACCCGCAAAAAACAGGCGTATGGGTGTGACATTACCCACTTCAGTGCTG  1147

Qy   961  TTTATCCGCGGACACGGATACTAATCTGCAAAATCTCGGCGGGGCCAGGTGGCTCAACTGG  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1148  TTTATCCGCGGACACGGATACTAATCTGCAAAATCTCGGCGGGGCCAGGTGGCTCAACTGG  1207

Qy  1021  ACGCTTCCCGGTCAGCCGGATAACACGCCGGCCAGGTGGTGAACTGGTGTGTTTGAACGCTGG  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1208  ACGCTTCCCGGTCAGCCGGATAACACGCCGGCCAGGTGGTGAACTGGTGTGTTTGAACGCTGG  1267

Qy  1081  CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGTTTCGCTCGTCTCTTCCAGACTTTACAG  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1268  CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGTTTCGCTCGTCTCTTCCAGACTTTACAG  1327

Qy  1141  CAGATGCGTGATTAAAAACGGCCGCTGTCATTAAATAACGCCGCCCGGAGGAGGTGAAAACTGACC  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1328  CAGATGCGTGATTAAAAACGGCCGCTGTCATTAAATAACGCCGCCCGGAGGAGGTGAAAACTGACC  1387

Qy  1201  CTGGCAGGATGTGAAAGAGGCGAAATGCGCAGTGTGTTCGTTGGGCAGGTTTTACGGCAA  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1388  CTGGCAGGATGTGAAAGAGGCGAAATGCGCAGTGTGTTCGTTGGGCAGGTTTTACGGCAA  1447

Qy  1261  ATCGTGAATGAAGGCACCCATACCGGCGGTGCAGTTTGTAAT  1300
          ||||||||||||||||||||||||||||||||||||||||||
Db  1448  ATCGTGAATGAAGGCACCCATACCGGCGGTGCAGTTTGTAAT  1487
```

RESULT 7
US-10-933-115-5
; Sequence 5, Application US/10933115
; Publication No. US20050281792A1
; GENERAL INFORMATION:
;   APPLICANT: Short, Jay M.
;   APPLICANT: Kretz, Keith A.
;   APPLICANT: Gray, Kevin A.
;   APPLICANT: Barton, Nelson Robert
;   APPLICANT: Garrett, James B.
;   APPLICANT: O' Donoghue, Eileen
;   APPLICANT: Mathur, Eric J.
;   TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
;   TITLE OF INVENTION: THEREOF
;   FILE REFERENCE: 564462001822
;   CURRENT APPLICATION NUMBER: US/10/933,115
;   CURRENT FILING DATE: 2004-09-01
;   PRIOR APPLICATION NUMBER: US/09/866,379
;   PRIOR FILING DATE: 2001-05-24
;   PRIOR APPLICATION NUMBER: US 09/580,515
;   PRIOR FILING DATE: 2000-05-25
;   PRIOR APPLICATION NUMBER: US 09/318,528
;   PRIOR FILING DATE: 1999-05-25
;   PRIOR APPLICATION NUMBER: US 09/291,931
;   PRIOR FILING DATE: 1999-04-13
;   PRIOR APPLICATION NUMBER: US 09/259,214
;   PRIOR FILING DATE: 1999-03-01
;   PRIOR APPLICATION NUMBER: US 08/910,798
;   PRIOR FILING DATE: 1997-08-13
;   NUMBER OF SEQ ID NOS: 10
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 1901
;   TYPE: DNA
;   ORGANISM: Escherichia coli
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 403
;   OTHER INFORMATION: n = A,T,C or G
US-10-933-115-5

Query Match   96.9%;   Score 1267;   DB 10;   Length 1901;
Best Local Similarity 98.4%;   Pred. No. 0;
Matches 1279; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
Qy    1  ATGAAAGCGGATCTTAAATCCCATATTTTTAATCTCTTCTTCGTATTCCGTTAACCCGCAATCTGCA   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  188  ATGAAAGCGGATCTTAAATCCCATATTTTTATCTCTTCTTCGTATTCCGTTAACCCGCAATCTGCA  247

Qy   61  TTCGCTCAGAGTGAGCCGGAGCTGAAAGCTGGAAAAGTGGTGTTCAGTGTCGTCGTCATGGT  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  248  TTCGCTCAGAGTGAGCCGGAGCTGAAAGCTGGAAAAGTGGTGTTCAGTGTCGTCGTCATGGT  307

Qy  121  GTGGGTGCTTCCAACCAAGCCACCGGAGCCGAACGCTGGCTGGAAAGCAGGATGTCAGTGGGCCA  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  308  GTGGGTGCTTCCAACCAAGCCACCGGAGCCGAACGCTGGCTGGAAAGCAGGATGTCAGTGGGCCA  367

Qy  181  ACCTGGCCGGTAAAAACTGGGTGAGCTGACAGCCGACCGGGCGGTGGGTAGGTAAATCGCCTATCTC  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  368  ACCTGGCCGGTAAAAACTGGGTGACGTGACAGCCGACCGGGCGGTGGGTAGGTAAATCGCCTATCTC  427

Qy  241  GGACATTACTGGGCGTCAGCGGTCTCGGTAGCCGCCGGATTGCTGCCTAAAAGTGGCTGCCCG  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  428  GGACATTACCAACCGCCAGCCGTCTCGGTAGCCGGATTGCTGGGCCGGAAAAAAGGGCTGCCCG  487

Qy  301  CAGTCTGGTTCAGGTTCGGCGATTATTGCTGGTTCGGCTAATTTCCGGCAATCAAACTTGTGC  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  488  CAGTCTGGTTCAGGTTCGGCGATTATTGCTGATTGTCGACGGACGGTACCGGTAAAAACAGGCGAA  547

Qy  361  GCCTTGCGCGCGGGGTCGGGCACCTGACTGTGCAATAACCGTAACATACCCAGGCAGATACG  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  548  GCCTTGCGCGCGGGGTCGGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG  607

Qy  421  TCCAGTCCCGATCCGTTATTAATCCTTAAAAAACTGGCGTTGCCCGG  480
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  608  TCCAGTCCCGATCCGTTATTTAATCCCTCTAAAAAACTGGCGTTGCCAACTGGATAAACGGCG  667

Qy  481  AAGGTGACTGACCGCGATCCTTCGGAGAGGGCCAGGAGGGGTCAAATTGCTGGACATTTACCGGCAT  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  668  AAGGTGACTGACCGCGATCCTTCGGAGAGGGCCAGGAGGGGTCAAATTGCTGGACATTTACCGGCAT  727

Qy  541  TATCAAACGGCGTTTCGCGAACTGGAACGGGGTGCTTAATTTCCGGCAATCAAACTTGTGC  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  728  CGGCGAACGGGCGTTTCGCGAACTGGAACGGGGTGCTTAATTTCCGGCAATCAAACTTGTGC  787

Qy  601  CTTAAACGTGAGAAACAGGACGAAAGCGTGTTCATTAACGGCAGGCGCATTAACCATCGGAACTC  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  788  CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCATTAACGGCAGGCGCATTAACCATCGGAACTC  847

Qy  661  AAGGTGAGGCGGCCGACTGTGTCTCGAGAGGGCGGGGTAAGCCTTCGCCATCAATCTGACG  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  848  AAGGTGACGCGCCGACAAGTGTCCTAACCGGTGCGGGGTAAGCCTTCGCCATCAATCTGACG  907

Qy  721  GAGAGTATTTCTTGCAACGAAGCACAGGGAATCCGGAGCCGGGGGGAAGGAGGATCACC  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  908  GAGAATATTTCTCCTGCAACAGCACAGGGAATGCCGGAGCCGGGGGGAAGGAGGATCACC  967

Qy  781  GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTGCTACA  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  968  GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTATTTGCTACA  1027

Qy  841  CGCACGCGCCAGAGGTTGCCCGCGACCAGCGCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1028  CGCACGCGCAGAGGTTGCCCGCGACCAGCGCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG  1087

Qy  901  TTGACGACGCCCATCCACCCGCAAAAAACAGGCGTATGGGTGTGACATTACCCACTTCAGTGCTG  960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1088  TTGACGACGCCCATCCACCCGCAAAAAACAGGCGTATGGGTGTGACATTACCCACTTCAGTGCTG  1147
```

RESULT 8

US-09-777-566A-1

; Sequence 1, Application US/09777566A
; Patent No. US20010055788A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-6
; CURRENT APPLICATION NUMBER: US/09/777,566A
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.0
;
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1320)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n is any nucleotide

US-09-777-566A-1

Query Match   96.7%; Score 1265; DB 3; Length 1323;
Best Local Similarity   98.0%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy   1 ATGAAAGCGATCTTAATCCCCATTTTTATTCTTCTTCGATTTCGTTAACCCGCCAATCTGCA 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAGCGATCTTAATCCCCATTTTTATTCTTCTTCGATTTCGTTAACCCGCCAATCTGCA 60

Qy   61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCATGGGT 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGTGGTGATTGTCAGTCATGGGT 120

Qy   121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGAGTGTCACCCCGAGCGATGCCA 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGAGTGTCACCCCGAGCGATGCCA 180

RESULT 9
US-09-866-379-1
; Sequence 1, Application US/09866379
; Patent No. US2002013675A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
;
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n is any nucleotide
; NAME/KEY: CDS
; LOCATION: (1)..(1323)
; OTHER INFORMATION:
US-09-866-379-1

Query Match   96.7%; Score 1265; DB 3; Length 1323;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```
Qy      1 ATGAAAGGGATCTTAATCCACATTTTATCTCTTCGATTCCGTTAACCCGCAAATCTGCA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAGGGATCTTAATCCACATTTTATCTCTTCGATTCCGTTAACCCGCAAATCTGCA 60

Qy     61 TTCGCTCAGAGTGAGCCGGGAGCTGAAGCTGAAGTGTCAGTCGTCATGGT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCGCTCAGAGTGAGCCGGGAGCTGAAGCTGAAAGTGTCAGTCGTCATGGT 120

Qy    121 GTGCGTGCTCCAACCAAGGCCACGGCAACTGATCGATCCAGACGCATGGCCA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGCGTGCTCCAACCAAGGCCACGGCAACTGATCGATCCAGACGCATGGCCA 180

Qy    181 ACCTGGCCGGGTAAAACTGGGTTGGGTTGGCTGCGCCG 240
            ||||||||||||||||||||||||||||||||||||||
Db    181 ACCTGGCCGGGTAAAACTGGGTTGGGTTGGCTGCGCCG 240

Qy    241 GGACATTACCAACGCCAGCGTCTGGTAGCCGAGCGGATTGCTGCGGCAAAAAACAGGGCAA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGACATTACCAACGCCAGCGTCTGGTAGCCGAGCGGATTGCTGCGGCAAAAAACAGGGCAA 300

Qy    301 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTGACGAGGCGATTACCCGTAAAAACAGGGCGAA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTGACGAGGCGATTACCCGTAAAAACAGGGCGAA 360
```

```
Db    301 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTGACGAGGCGATTACCCGTAAAAACAGGGCGAA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    361 GCCTTCGCCGCCCGCCCGGCTGGCGCACCTGGCATTATTGCTGAAGTCTACAT ACCCAGGCGAGATACG 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCTTCGCCGCCCGCCCGGCTGGCGCACCTGGCATTATTGCTGAAGTCTACATACCCAGGCGAGATACG 420

Qy    421 TCCAGTCCCCGGATCCGGTTATTTAATCCTCTAAAAAACTGGGGTTTGCCAACTGGATAACGCG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCCAGTCCCGGATCCGGTTATTTAATCCTCTAAAAAACTGGGGTTTGCCAACTGGATAACGCG 480

Qy    481 AACGTGACTGAACGGCGATCCTCGAGGAGGGGCAGGAGGGGTCAATTGCTGACTTTACCGGGCAT 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AACGTGACTGAACGGCGATCCTCGAGGAGGGGCAGGAGGGGTCAATTGCTGACTTTACCGGGCAT 540

Qy    541 CGGGCAAAACGGCGTTTCGGCGAAACTGGGAAACGGGTGCTTAAATTTTCCGGCAAACTTGTGC 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CGGGCAAAACGGCGTTTCGGCGAAACTGGGAAACGGGTGCTTAAATTTTCCGGCAAACTTGTGC 600

Qy    601 CTTAAACGTGAGAAAACAAGAGCAGAAGAGAAAGCTGTTCAATTAACGGCAGGCGATTACCATCCGAACTC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTTAAACGTGAGAAAACAAGAGCAGAAGAGAAAGCTGTTCAATTAACGGCAGGCGATTACCATCCGAACTC 660

Qy    661 AAGGTGAGGCGCGACTGTGTCTCATTAACCGGTGAAGCCTCGCATCAAATGCTGAGG 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AAGGTGAGGCGCGACTGTGTCTCATTAACCGGTGAAGCCTCGCATCAAATGCTGAGG 720

Qy    721 GAGATATTTCTTCCTGCAACAAGCACAGGGAATGCCAGACGCCGGGGTGGGGAAGGATCACC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GAGATATTTCTTCCTGCAACAAGCACAGGGAATGCCAGACGCCGGGGTGGGGAAGGATCACC 780

Qy    781 GATTCACCACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCGCAATTTGATTTGCTACAA 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GATTCACCACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCGCAATTTTATTTGCTACAA 840

Qy    841 CGCACGCCAGAGAGGTTGCCCGGCGCGCCAGCCCCGTTATTAGATTGATCAAGACAGCG 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CGCACGCCAGAGAGGTTGCCCGGCGCGCCAGCCCCGTTATTAGATTGATCAAGACAGCG 900

Qy    901 TTGACGCCCCATCCACCGCAAAAAACAGGGTTATGGTGTTGACATTACCCACTTCAGTTACTG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTGACGCCCCATCCACCGCAAAAAACAGGGTTATGGTGTTGACATTACCCACTTCAGTTACTG 960

Qy    961 TTTATTGCCGGGACACGGATACTAATCTGGCAAAATCTTCGGCGGGGCACTGGAGCTCAAACTGG 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTTATTGCCGGGACACGGATACTAATCTGGCAAAATCTTCGGCGGGGCACTGGAGCTCAAACTGG 1020

Qy   1021 ACGCTTCCCGGTCAGCCGGATAACACGCGCCAGTGGTGGTGAACTGGTGTTTGTTTGAACGGCTGG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGCTTCCCGGTCAGCCGGATAACACGCGCCAGTGGTGGTGAACTGGTGTTTGTTTGAACGGCTGG 1080

Qy   1081 CGTCGGGCTAAGCGATAACAGCCAGTGGATTCAGTTTCGCTGGTCTTCCAGACTTTACAG 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CGTCGGGCTAAGCGATAACAGCCAGTGGATTCAGTTTCGCTGGTCTTCCAGACTTTACAG 1140

Qy   1141 CAGATGCGTGATAAAAACGGCGCTGTCATTAAATAACGGCCGGAGAGGTGAAAACTGACC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAGATGCGTGATAAAAACGGCGCTGTCATTAAATAACGGCCGGAGAGGTGAAAACTGACC 1200

Qy   1201 CTGGCAGGATGTGAAAGAGGGAAATGCGCAGGGCCATGTGTTGCGTTTTTACGCAA 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGCAGGATGTGAAAGAGGGAAATGCGCAGGGCCATGTGTTGCGTTTTTACGCAA 1260

Qy   1261 ATCGTGAATGAAGCACGCAATACCCGGCGTGCAGTTTGAGATCTCATC 1306
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAGCACGCAATACCCGGCGTGCAGTTTGAGATCTCATC 1306
```

RESULT 10
US-10-034-985-1
; Sequence 1, Application US/10034985
; Publication No. US20030049815A1
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith

```
Qy   1261 ATCGTGAATGAAGCACGCAATACCCGGCGTGCAGTTTGAGATCTCATC 1306
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAGCACGCAATACCCGGCGTGCAGTTTGAGATCTCATC 1306
```

```
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/034,985
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
;
US-10-034-985-1
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G

Query Match      96.7%;  Score 1265;  DB 6;  Length 1323;
Best Local Similarity  98.0%;  Pred. No. 0;
Matches 1280;  Conservative  0;  Mismatches  26;  Indels  0;  Gaps  0;
```

```
Qy      1 ATGAAAGCGGATCTTAATCCCATTTTTATTCTTCTTCTGATTTACCGTGCCCAATCTGCA   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAGCGGATCTTAATCCCATTTTTATTCTTCTTCTGATTTACCGTGCCCAATCTGCA   60

Qy     61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGAAAGTGTGGTGGATTGTTCAGTCGTCGGT  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGAAAGTGTGGTGGATTGTTCAGTCGTCATGGT  120

Qy    121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGAGTGTCACCCCGAGACGCATGGCCA  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGAGTGTCACCCCGAGACGCATGGCCA  180

Qy    181 ACCTGGCCGGTAAAACTGGGGTGAGCTGCACACACCGGCAGGCTAATCGGCCTATCTC  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACCTGGCCGGTAAAACTGGGGTGAGCTGCACACACCGGCAGGCTAATCGGCCTATCTC  240

Qy    241 GGACAATTACTGGGCGTCAGCGGTCAGGCGTAGCCCGGCGAGAAAAGGGCTGCCCG  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGACAATTACCAAACGCCAGCGGTCTGTGTAGCCCGGCGAGAAAAGGGCTGCCCG  300

Qy    301 CAGTCGTGGTCAGGTCGCGGATTATTGCTGATGTGTGGACCGAGCGGTACCCGGTACCGGAA  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGTCGTGGTCAGGTCGCGGATTATTGCTGATGTGTGGACCGAGCGGTACCCGGTACCGGAA  360

Qy    361 GCCTTCGCCCGCGGCGGCTGGCACCTGACTGTGCAATAACCGTACCATACCCAGGCACGATACG  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCTTCGCCCGCGGCGGCTGGCACCTGACTGTGCAATAACCGTACCATACCCAGGCACGATACG  420

Qy    421 TCCAGTCCCCGATCCGTTATTTAATCCTCTAAAAACTGGGATAACGGCG  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCCAGTCCCCGATCCGTTATTTAATCCTCTAAAAACTGGGATAACGGCG  480

Qy    481 AACGTGACTGACGCGGATCCTCGAGAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AACGTGACTGACGCGGATCCTCAGCAGGGGTCAATTGCTGACTTTACCGGGCAT  540

Qy    541 TATCAAAACGCGGTTTCGCGAACTGGAACGGGGTGCTTAATTTTTCCGCAATACAAAACTTGTGC  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CGGCAAACGGCGGTTTCGCGAACTGGAACGGGGTGCTTAATTTTTCCGCAATACAAAACTTGTGC  600

Qy    601 CTTAAACGTGAGAAAACAGGAGACGAAAGCTGTTCATTAACGCAGGCAATTACCATCGGAACTC  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTTAAACGTGAGAAAACAGGAGACGAAAGCTGTTCATTAACGCAGGCAATTACCATCGGAACTC  660
```

```
Qy    661 AAGGTGAGCGGCCGACGTGTTCTCATTAACCGGTGCGGTGCGTAAGGCTTCGGCATCAATGCTTGACG  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AAGGTGAGCGGCCGACGAATGTCTTCATTAACCGGTGCGGTGCGTAAGGCTTCGGCATCAATGCTTGACG  720

Qy    721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGGTGGGGAGGGATCACC  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGGTGGGGAGGGATCACC  780

Qy    781 GATTCACACCAGTGGAACACCCTTGCTTAAGTTTGCATAAACGGGGAGGGGGTGCTACAA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GATTCACACCAGTGGAACACCCTTGCTTAAGTTTGCATAAACGGGCAATTTGATTTGCTACAA  840

Qy    841 CGCACGCCAGAGGTTGCCCGGCAGCCCGCCGTTATTAGATTTGATCAAGACACAGCG  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CGCACGCCAGAGGTTGCCCGGCAGCCCGCCCGTTATTAGATTTGATCATGGCAGCG  900

Qy    901 TTGACGCGCCCATCCACCGCAAAAAACAGGCGTARGGTGTGAGGTGAACTACTG  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTGACGCGCCCATCCACCGCAAAAAACAGGCGTATGGTGTGAGGTGACATTACCCACTTCAGTACTG  960

Qy    961 TTTATCGCCGGACACGATACTAATCCTGGCAAAATCTCGGGCAACGGCCACTGGAGCTCAACTGG  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTTATTGCCCGGACACGATACTAATCCTGGCAAAATCTCGGGCGGCCACTGGAGCTCAACTGG  1020

Qy   1021 ACGGCTTCCCGGTCAGCGCGGATAAACACCCGCCGCCAGGTGGTGAACTGGTGTTTGAACCGCTGG  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGGCTTCCCGGTCAGCGCGGATAAACACCCGCCGCCAGGTGGTGAACTGGTGTTTGAACCGCTGG  1080

Qy   1081 CGTCGGCTAAGCGGCTAACAGCCAGTGGATTCAGGTTTCGTCGTTCTTCCAGACTTTACAG  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CGTCGGCTAAGCGGCTAACAGCCAGTGGATTCAGGTTTCGTCGTTCTTCCAGACTTTACAG  1140

Qy   1141 CAGATCCGTTGATAAAACGCCCCTGTCATTAAATAACCCGCCGGAGGTGAAAACTGACC  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAGATGCGGTGATAAAACGCCGTCATTAAATAACCCGCCGGAGGTGAAAACTGACC  1200

Qy   1201 CTGGCGCAGGATGTGAAGCGGAAATGGGCAGCGCATCGTTCGTTTGGCAGGGTTTTACGGCAA  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGCAGGATGTGAAGAGCGAAATGGGCAGCGCATCGTTCGTTTGGCAGGGTTTTACGGCAA  1260

Qy   1261 ATCGTGAATGAAGCACGCCATACCCGGCGGCTGCAGTTTGAGATCTCATC  1306
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAGCACGCCATACCCGGCGGCTGCAGTTTGAGATCTCATC  1306
```

```
RESULT 11
US-10-430-356-1
; Sequence 1, Application US/10430356
; Publication No. US20030232041A1
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/430,356
; CURRENT FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
;
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
```

```
;  NAME/KEY: CDS
;  LOCATION: (1)...(1320)
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (1)..(1323)
;  OTHER INFORMATION: n = A,T,C or G
US-10-430-356-1

Query Match        96.7%;  Score 1265;  DB 7;  Length 1323;
Best Local Similarity   98.0%;  Pred. No. 0;
Matches 1280;  Conservative  0;  Mismatches  26;  Indels  0;  Gaps  0;

Qy      1 ATGAAAACGATCTTAATCCCATTTTTACTCTCTGATTCCGATTAACCCCGAAATCTGCA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAAGGATCTTAATCCCATTTTTACTCTCTGATTCCGATTAACCCCGAAATCTGCA 60

Qy     61 TTCGCTCGAGAGTGAGCCCGGAGCTGAAGCTGGAAAGTGAGTGAGTGCAGTCGTTCATGGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCGCTCAGAGTGAGCCCGGAGCTGAAGCTGGAAAGTGTGGGTGAAAGTGTCAGTCCATGGT 120

Qy    121 GTGCGTGCTCCAACCAAGGCCACCAGCCAACTGATGCAGAGATCACCCCAGACACGCATGGCCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGCGTGCTCCAACCAAGGCCACCAGCCAACTGATGCAGCATGTCACCCCCAGACGCCA 180

Qy    181 ACCTGGCCGGGTAAAACTGGGTTGGGTGAGCTCACACCCCGCCGCGGCTAATCGCCTACTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACCTGGCCGGGTAAAACTGGGTTGGGTGAGCTGACACCCCGCCGCGGCTAATCGCCTACTC 240

Qy    241 GGACATTACTAGCCGGTCAGGTCGGCGTAGCCGACGGATTGCTGCCTAAAGTGGCTGCCCG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGACATTACCAACGCCAGCGTCTGGGTAGCGACGGATTGCTGGCGAAAAAAGGGCTGCCCG 300

Qy    301 CAGTCTCGGTCAGGTCGGCGGATTAATTGCTGATGTCGACGAGGCGTACCCGTAAAACAGGCGAA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGTCTCGGTCAGGTCGGCGGATTAATTGCTGATGTCGACGAGGCGTACCCGTAAAACAGGCGAA 360

Qy    361 GCCTTCGCCGCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCTTCGCCGCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420

Qy    421 TCCAGTCCCGGATCCGGTTATTTTAATCCTCTAAAAACTGGGCTGGTTTGCCAACTGGATAACGCG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCCAGTCCCGGATCCGGTTATTTTAATCCTCTAAAAACTGGGCTGGTTTGCCAACTGGATAACGCG 480

Qy    481 AACGTGACTGAGCGGATCCTCAGAGGGCGCAGAGGGTCAATGCTGACTTTTACCGGGGCAT 540
          ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db    481 AACGTGACTGAGCGGATCCTCAGCAGGGCAGAGGGTCAATGCTGACTTTTACCGGGGCAT 540

Qy    541 TATCAAAACGGGGTTTGGCGGAACTGGGAACGGGTGCTTAATTTTTCCGGCAAATCAAACTTGTGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CGGCAAAACGGGCGTTTCCCGACTGGAACGGGTGCTTAATTTTTCCGGCAAATCAAACTTGTGC 600

Qy    601 CTTAAACGTGAGAAAACAGGAACGGAAAGCTGTTCATTAACGAGCAGGCCATTAACCATTCGGGAACTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTTAAACGTGAGAAAACAGGAACGGAAAGCTGTTCATTAACGAGCAGGCCATTAACCATTCGGGAACTC 660

Qy    661 AAGGTGCAGCGCCGACTGTGTCTCATTAACCGGTGCGGTAAGCCTCGCATCCAATGCTGACG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AAGGTGAGCGGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 720

Qy    721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGTGGGAAGGATCACC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GAGATATTTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGTGGGAAGGATCACC 780

Qy    781 GATTCACCCAGTGGAACACCTTGCTAAGTTGCATAACGCGGCCAATTTATTTGCTACAA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GATTCACCCAGTGGAACACCTTGCTAAGTTGCATAACGCGGCCAATTTATTTGCTACA 840

Qy    841 CGCACGCCAGAGGTTGCCCGCCAGCCGCCACCCGGTTATTGGGATTTGATCATGGCAGGCG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CGCACGCCAGAGGTTGCCCGGCAGCCGCCACCCGGTTATTGGGATTTGATCATGGCAGGCG 900

Qy    901 TTGACGCCCCATCCACCGCGAAAAACAGGCGGTATGGGTGGTGACATTACCCACTTCAGGTCCTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTGACGCCCCATCCACCGCGAAAAACAGGCGGTATGGGTGGTGACATTACCCACTTCAGGTCCTG 960

Qy    961 TTTATTGCCCGGACACGGATACTAATCTGGCAAAATCTCGGCACTCGGCG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTTATTGCCCGGACACGGATACTAATCTGGCAAAATCTCGGCACTCGGCG 1020

Qy   1021 ACGCTTCCCGGTCAGCCGGATAACGACGCCGGCCCAGGTGGGTGAACTGGTGTGAACGCTGG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGCTTCCCGGTCAGCCGGATAACGACGCCGGCCCAGGTGGGTGAACTGGTGTGAACGCTGG 1080

Qy   1081 CGTCCGGCTAAGCGACAGCCAGTGGATCAGGTTTCAGGTTTCCGTCGTTCCAGACTTTACAG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CGTCCGGCTAAGCGACAGCCAGTGGATCAGGTTTCAGGTTTCCGTCGTTCCAGACTTTACAG 1140

Qy   1141 CAGATGCGTGGATAAAACGCGCTGTCATTAAAAATACGCCCGCCGGGGCATGTGTTCGGGCCGG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAGATGCGTGATAAAACGCGCTGTCATTAAAAATACGCCCGCCGGGGCATGTGTTCGGGCCGG 1200

Qy   1201 CTGGCCAGGATGTGAAGAGGAGGCGAGGGCATGTGTTCGTTCGTTTGGCAGGTTTTTACGGCAA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGCCAGGATGTGAAGAGGAGGCGAAGGCAGGGCATGTGTTCGTTCGTTTGGCAGGTTTTTACGGCAA 1260

Qy   1261 ATCGTGAATGAAAGCACGCATACCGGCGTGCGAGTTTGAGACTCTCATC 1306
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAAGCACGCATACCGGCGTGCGAGTTTGAGACTCTCATC 1306
```

RESULT 12
US-10-601-319-1

Query Match   96.7%; Score 1265; DB 8; Length 1323;

Left column

Best Local Similarity 98.0%; Pred. No. 0; Mismatches 26; Indels 0; Gaps 0;
Matches 1280; Conservative 0;

```
Qy      1 ATGAAAGCGATCTTAATCCCATTTTTATCTTCTTGATTCCGTTAACCCCGCAATCTGCA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTCTTGATTCCGTTAACCCCGCAATCTGCA 60

Qy     61 TTCGCTCAGAGTGAGCCGGACTGAAGCTGAAAGTGGTGATTGTCAGTCGTCGTCATGGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCGCTCAGAGTGAGCCGGACTGAAGCTGAAAGTGGTGATTGTCAGTCGTCGTCATGGT 120

Qy    121 GTGCCGTTGCTCCAACCAAGGCCACGCAACTGATGCACCCCAGACGGCCATGGCCA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGCCGTTGCTCCAACCAAGGCCACGCAACTGATGCACCCCAGACGGCCATGGCCA 180

Qy    181 ACCTGGCCGGGTAAAAACTGGGTTGGCTAATGCCTATCTC 240
          |||||||||||||||||||||||||||||||||||||||
Db    181 ACCTGGCCGGGTAAAAACTGGGTTGGCTGACACCCCGCGGGCTGGCTAATGCCTATCTC 240

Qy    241 GGACATTACTGGCGTCAGCGTCGTCGGTAGCCGACGGATTGCTGCCTAAATGTGGGCTGCCCG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGACATTACCAACGCCAGCGGTCTGGGTAGCCGACGGATTGCTGCCTAAATAACAGGGCTGCCCG 300

Qy    301 CAGTCTTCGGTCAGGTCGGCGATTATTGCTGATTGCTGAGTCGACGGCGTACCCGTAAAAACAGGCGAA 360
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGTCTTCGGTCAGGTCGGCGATTATTGCTGATTGCTGAGTCGACGGCGTACCCGTAAAAACAGGCGAA 360

Qy    361 GCCTTCGCCGCGGCGTTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCCAGATACG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCTTCGCCGCGGCGTTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCCAGATACG 420

Qy    421 TCCAGTCCCGATCCGTTGTTATTTAATCCTCTAAAAACTGGATAACGCG 480
          |||||||||||||||||||||||||||||||||||||||||
Db    421 TCCAGTCCCGATCCGTTTATTTAATCCTCTAAAAACTGGATAACGCG 480

Qy    481 AACGTGACTGACGCGATCCTCGAGAGGGCAGGAGGGTCAATTGCTGAACTTTACCGGGCAT 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AACGTGACTGACGCGATCCTCGAGCAGGGCAGGAGGGTCAATTGCTGAACTTTACCGGGCAT 540

Qy    541 TATCAAACGCGGTTTCGCGAACTGGAACGGGGTGCTTAATTTTCCGCAAATCAAACTTGTGC 600
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CGGCAAACGCGGGTTTCGCGAACTGGAACGGGGTGCTTAATTTTCCGCAAATCAAACTTGTGC 600

Qy    601 CTTAAACGTGAGAAAACAGGAGAACTGTTCATTAACGACGCATTACCATCGGGAACTC 660
          ||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTTAAACGTGAGAAAACAGGAGAACTGTTCATTAACGACGCATTACCATCGGGAACTC 660

Qy    661 AAGGTGAGCGGCCGACTGTGTCTCATTAACCGGTGCGGGTAAGCCTCGGCATCAAGCTGACG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AAGGTGAGCGGCCGACAATGTCTCATTAACCGGTGCGGGTAAGCCTCGGCATCAATGCTGACG 720

Qy    721 GAGAGATATTTCTCCTGCAAACAAGCACACAGGGGAATGCCGGACGGCCGGGTGGGGAAGGATCACC 780
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GAGAGATATTTCTCCTGCAAACAAGCACACAGGGGAATGCCGGACGGCCGGGTGGGGAAGGATCACC 780

Qy    781 GATTCACACCAGTGGAACACCTTGCTAAGTTGCATAACGGCAATTTGATTTGCTACAA 840
          |||||||||||||||||||||||||||||||||||||||||||||
Db    781 GATTCACACCAGTGGAACACCTTGCTAAGTTGCATAACGGCAATTTTATTTGCTACAA 840

Qy    841 CGCACGCGCCAGAGGGTTGCCCGCGCGCCACCCGTTATTGGATTTGATCAAGACAGCG 900
          ||||||||||||||||||||||||||||||||||||||||||||
Db    841 CGCACGCCCAGAGGTTGCCCGGCCGCCACCCGTTATTGGATTTGATCAATGCAGAGCG 900

Qy    901 TTGACGCCCCATCCACCGTATGGGTGTGACATTACCCACTTCAGTGCTG 960
          |||||||||||||||||||||||||||||||||||||||||
Db    901 TTGACGCCCATCACCGGTATGGGTGTGACATTACCCACTTCAGTGCTG 960

Qy    961 TTTATTGCCGGACAACGATTAATCTGGGCGACGCACTGGGACTTCAACTGG 1020
          |||||||||||||||||||||||||||||||||||||||||
Db    961 TTTATTGCCGGACAACGATTAATCTGGGCGACGCACTGGGGACTTCAACTGG 1020

Qy   1021 ACGCTTCCCGGTCAGCGGGATAACACACCCGCCAGGTGGTGAACTGGTGTTCGAACGCTGG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGCTTCCCGGTCAGCGGGATAACACACCCGCCAGGTGGTGAACTGGTGTTCCAGACTTTACAG 1080

Qy   1081 CGTCGGCTAAGCCATAACAGCCAGTGGGATTCAGGTTTCGCTGGGTCTTCCAGACTTTACAG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CGTCGGCTAAGCCATAACAGCCAGTGGGATTCAGGTTTCGCTGGGTCTTCCAGACTTTACAG 1140

Qy   1141 CAGATACGTGATAAAAACGCCCGCCTGTCATTAAATAATACCGGGCGGAGAGGTGAAACTGACC 1200
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAGATACGTGATAAAAACGCCCGCCTGTCATTAAATAATACCGGTGGAGAGGTGAAACTGACC 1200

Qy   1201 CTGGCACGGATGGTGAAGCGAAATGGGCCAGCATTGTTGGCAGGTTTTTACGCAA 1260
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGCACGGATGGTGAAGAGCGAAATGGGCCAGCATTGTTGTTCGTTTGGCAGGTTTTTACGCAA 1260

Qy   1261 ATCGTGAATGAAGCACGCCATACCGGGCGGCTGCAGTTTGAGATCTCATC 1306
          |||||||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAGCACGCCATACCGGGCGGCTGCAGTTTGAGATCTCATC 1306
```

Right column

RESULT 13
US-10-933-115-1

; Sequence 1, Application US/10933115
; Publication No. US20050281792A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 564462001822
; CURRENT APPLICATION NUMBER: US/10/933,115
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 216
; OTHER INFORMATION: n = A,T,C or G
US-10-933-115-1

Query Match 96.7%; Score 1265; DB 10; Length 1323;
Best Local Similarity 98.0%; Pred. No. 0; Mismatches 26; Indels 0; Gaps 0;
Matches 1280; Conservative 0;

```
Qy      1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTTGATTCCGTTAACCCCGCAATCTGCA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTCTTGATTCCGTTAACCCCGCAATCTGCA 60

Qy     61 TTCGCTCAGAGTGAGCCGGACTGAAGCTGAAAGTGGTGATTGTCAGTCGTCGTCATGGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCGCTCAGAGTGAGCCGGACTGAAGCTGAAAGTGGTGATTGTCAGTCGTCGTCATGGT 120
```

```
Qy      61 TTCGCTCAGAGTGAGCCCGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTTCATGGT 120

Qy     121 GTGGCGTGCTCCAACCAAGGCCACCGCAGGATGCAGAACGTCACCCCAGACGCATGGTCTCATGGT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GTGGCGTGCTCCAACCAAGGCCACCGCAGGATGCAGAACGTCACCCCAGACGCATGGCCA 180

Qy     181 ACCTGGCGCCGGTAAAACTGGGTGAGCTGAGCATCGACACCCCGCCGGTTCATCGGCCTATCTC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ACCTGGCGCCGGTAAAACTGGGTGAGCTGAGCATCGACACCCCGCCGGTTCATCGGCCTATCTC 240

Qy     241 GGACATTACTACTGGGCTCAGCGTTCAGGCCGACCGGAATTGCTGCCTAAAATGTGGCTGCCCG 300
                                        ||||||
Db     241 GGACATTACCAGCACGCGTCTGGTTAGCCGTACCGGACGGATTGCTGCGGAAAAAGGGCTGCCCG 300

Qy     301 CAGTCGTTGGGTCAGGTGCGCGATTATTGCTGAGTGTCGACGACGTACCCGGTAAAAACAGGCGGAA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CAGTCGTTGGTCAGGTTCGGCGGATTATTGCTGAGTGTCGACGACGTACCCGGTAAAAACAGGGCAA 360

Qy     361 GCCTTCGCCGCCGCGGGCTCGGCCACCTGACTTGTGCAATAAGTCGACGTACCCGGTAAAAACAGGGCAA 420
           ||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GCCTTCGCCGCCGCGGGCTCGGCCACCTGACTTGTGCAATAAGTCGACGTACCCGGTAAAAACAGGGCAA 420

Qy     421 TCCAGTCCCGATCCGTTGTTATTTAATCCTCTAAAAAACTGGGCTTTGGGCTTGCCAACTGGATAAGGCG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TCCAGTCCCGATCCGTTGTTATTTAATCCTCTAAAAAACTGGGGTGTTGGGCTTGCCAACTGGATAAGGCG 480

Qy     481 AACGTGACTGACCGGATCCTCGAGAGGGCGAGAGGGTCAATTGCTGACTTTACCGGGGCAT 540
                                        ||||||
Db     481 AACGTGACTGACCGGATCCTCAGCAGGGCGAGAGGGTCAATTGCTGACTTTACCGGGGCAT 540

Qy     541 TATCAAACAGGGTTTCGCGAACTGGGAACAGGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600
                                        
Db     541 CGGCAAACGGCCGTTTCGCGAACTGGGAACAGGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600

Qy     601 CTTAAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGGCATTACCGGAACTC 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 CTTAAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGGCATTACCGGAACTC 660

Qy     661 AAGGTGAGCGCGCCGACTCGTGTCTCCATTAACCGGTGCGGTTAAGCCTTCGCATCAATGCTGACG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AAGGTGAGCGCGCCGACAATGTCTCTCATTAACCGGTGCGGTTAAGCCTTCGCATCAATGCTGACG 720

Qy     721 GAGATATTTCTCCTGCAACACAGCCACACCGGGAATGCCGGAGGCTGGAGAAGGATCACC 780
           ||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GAGATATTTCTCCTGCAACAAGCAGGGAATGCCGGAGGCTGGAGAAGGATCACC 780

Qy     781 GATTCACACCAGTGGAACACCTTGCTCAAGTTGCATAAACGCGCAAATTTATTATTTGCTACAA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GATTCACACCAGTGGAACACCTTGCTAGTTGCATAAACGCGCAAATTTATTATTTGCTACAA 840

Qy     841 CGCACGCCAGAGGTTGCCCGCAGCCGCCACCCCGTTATTCGGATTTGATCAACGGCAGCG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CGCACGCCAGAGGTTGCCCGCAGCCGCCACCCCGTTATTCGGATTTGATCAACGGCAGCG 900

Qy     901 TTGACCGCCCCATCCACCGCAAAAACAGGCGTATGGGTTGACATTACCCACTTCAGTTACTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 TTGACCGCCCCATCCACCGCAAAAACAGGCGTATGGGTTGACATTACCCACTTCAGTTACTG 960

Qy     961 TTTTATCCCGGACACGATACTAATCTGGCAAAATCTCGGCGGCGGAGTCGAAGCTCAACTGG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 TTTATTGCCCGGACACGATACTAATCTGGCAAAATCTCGGCGGCGGAGTCGAAGCTCAACTGG 1020

Qy    1021 ACGCTTCCCGGTCAGCGGGATTAAACACCCGCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 ACGCTTCCCGGTCAGCGGGATTAAACACCCGCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1080

Qy    1081 CGTCGGCTAACGGATAAACAGCCAGTGGCATGGTTTCGCGTGGTTCGCTTCCAGACTTTACAG 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CGTCGGCTAAACGGATAAACAGCCAGTGGCATGGTTTCGCGTGGTTCGCTTCCAGACTTTACAG 1140

Qy    1141 CAGATGGGTGATAAAACGCCGCTGTGTCTCATTAAATACGCGGCGCCGGAGGTGAAAACTGACC 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 CAGATGGGTGATAAAACGCCGCTGTGTCTCATTAAATACGCGGCGCCGGAGGTGAAAACTGACC 1200
```

RESULT 14

US-10-282-122A-7167

; Sequence 7167, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 7167
; LENGTH: 1299
; TYPE: DNA
; ORGANISM: Escherichia coli

US-10-282-122A-7167

Query Match   96.6%; Score 1264; DB 8; Length 1299;
Best Local Similarity   98.5%; Pred. No. 0;
Matches 1276; Conservative   0; Mismatches 20; Indels   0; Gaps   0;

```
Qy       1 ATGAAAGCGGATCTTAATCCATTTTACTCTTCTGATTCCGTTAACCCCGCAATCTCGCA 60
           |||||||||  ||||||||||||||||  |||||  ||||||||||||||||||||||||
Db       1 ATGAAAGCGATCTTAATCCATTTTATCTCTTTGATTCCGTTAACCCCGCAATCTCGCA 60

Qy      61 TTCGCTCAGAGTGAGCCCGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTTCATGGT 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TTCGCTCAGAGTGAGCCCGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTTCATGGT 120

Qy     121 GTGGCGTGCTCCAACCAAGGCCACCGCAGGATGCAGAACGTCACCCCAGACGCATGGCCA 180
```

```
Qy    1201 CTGGCCAGGATGTGAAGAGCGAGGCGAAAATCCGCAGGGGCGCATGTTGTTCGGTTCGGCCAGGTTTTTACGGCAA 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 CTGGCCAGGATGTGAAGAGCGAGGCGAAAATCCGCAGGGGCGCATGTTGTTCGGTTCGGCCAGGTTTTTACGGCAA 1260

Qy    1261 ATCGTGAATGAAGCACGCCAATACCCGCGCGTGCAGTTTGAGATCTCATC 1306
           |||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 ATCGTGAATGAAGCACGCCAATACCCGCGCGTGCAGTTTGAGATCTCATC 1306
```

RESULT 15
US-09-866-379-6

; Sequence 6, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
;   APPLICANT: DIVERSA CORPORATION
;   APPLICANT: SHORT, Jay
;   APPLICANT: KRETZ, Keith
;   APPLICANT: GRAY, Kevin
;   APPLICANT: BARTON, Nelson
;   APPLICANT: GARRETT, James
;   APPLICANT: O'DONOGHUE, Eileen
;   TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
;   FILE REFERENCE: DIVER1370-7
;   CURRENT APPLICATION NUMBER: US/09/866,379
;   CURRENT APPLICATION NUMBER: DIVER1370-7
;   CURRENT FILING DATE: 2001-05-24
;   PRIOR APPLICATION NUMBER: US 09/580,515
;   PRIOR FILING DATE: 2000-05-25
;   PRIOR APPLICATION NUMBER: US 09/318,528
;   PRIOR FILING DATE: 1999-05-25
;   PRIOR APPLICATION NUMBER: US 09/291,931
;   PRIOR FILING DATE: 1999-04-13
;   PRIOR APPLICATION NUMBER: US 09/259,214
;   PRIOR FILING DATE: 1999-03-01
;   PRIOR APPLICATION NUMBER: US 08/910,798
;   PRIOR FILING DATE: 1997-08-13
;   NUMBER OF SEQ ID NOS: 10
;   SOFTWARE: Patentin version 3.1
; SEQ ID NO 6
;   LENGTH: 1901
;   TYPE: DNA
;   ORGANISM: Escherichia coli
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1901)
;   OTHER INFORMATION: n is any nucleotide
US-09-866-379-6

Query Match   96.6%;   Score 1263.8;   DB 3;   Length 1901;
Best Local Similarity   98.2%;   Pred. No. 0;   Mismatches   23;   Indels   0;   Gaps   0;
Matches 1277; Conservative   0;

Qy   1 ATGAAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAAATCTGCA   60
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 ATGAAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAAATCTGCA   247

Qy   61 TTCGCTCAGAGTGAGCCCGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCAGTCATGGT   120
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248 TTCGCTCAGAGTGAGCCCGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCAGTCATGGT   307

Qy   121 GTGCGTGTCTCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCGACCGCATGGCCA   180
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308 GTGCGTGTCTCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCGACCGCATGGCCA   367

Qy   181 ACCTGGCCGCGGTAAAACTGGGGTGAGCTGACACCGCGCGCGGGTGACGGTAATCGCCTATCTC   240
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368 ACCTGGCCGCGGTAAAACTGGGGTGAGCTGACACCGCGCGCGGGTGACGGTAATCGCCTATCTC   427

Qy   241 GGACATTACTGGCGTCAGCGTCTGGTAGCCGACGACGGATTGCTGCCTAAATGTGGCGCCCG   300
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 GGACATTACTGGCGTCAGCGTCTGGTAGCCGACGACGGATTGCTGCCTAAATGTGGCGCCCG   487

Qy   301 CAGTCTCGGTCAGGTCGCCGGATTATTGCTGCAGTGTCGACGATTATTGCTGCGACGTGTTTACGCCA   360
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488 CAGTCTCGGTCAGGTCGCCGGATTATTGCTGCAGTGTCGACGATTATTGCTGCGACGTGTTTACGCCA   547

Qy   1201 CTGGCAGGAGGATGTGAAGAGCGAAATGCCAGGCGTGCCAGTGTTG   1260
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAGCACCGCCATACCGGCCGTGCAGTTTG   1296

Qy   1261 ATCGTGAATGAAGCACCGCCATACCGGCCGTGCAGTTTG   1296
       |||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAGCACCGCCATACCGGCCGTGCAGTTTG   1296

```
Qy    361 GCCTTCGCCCGCCCGGCCCCTTGCGCACTGTGCAATAAACCGTACATTACCCAGGCAGATACG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548 GCCTTCGCCCGCCCGGCCCCTTGCGCACCTGCAATAACCGTACATTACCCAGGCAGATACG 607

Qy    421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAAACTGGGGTTTGCCAACTGGATAAACGCG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAAACTGGGGTTTGCCAACTGGATAAACGCG 667

Qy    481 AACGTGACTGACGGCGGATCCTTCGAGAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    668 AACGTGACTGACGGCGGATCCTTCGAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 727

Qy    541 TATCAAAACGGGGCGTTTCGGGAAACTGGGAACGGGGTCGCTTAATTTTCCGCAATCAAAACTTGTGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    728 CGGCAAACGGGGCGTTTCGGGAAACTGGGAACGGGGTCGCTTAATTTTCCGCAATCAAAACTTGTGC 787

Qy    601 CTTAAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGGCATTACCATCGGGAACTC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    788 CTTAAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGGCATTACCATCGGGAACTC 847

Qy    661 AAGGTGAGCGCGACTGTGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATCGCTGACG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    848 AAGGTGAGCGCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 907

Qy    721 GAGATATTTTCTCCTGCAACAAGCACAGGGAAATGCCGGAGCCGGGGTGGGGGAGGGATCACC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908 GAGATATTTTCTCCTGCAACAGGCACAGGGAAATGCCGGAGCCGGGGTGGGGAAGGGATCACC 967

Qy    781 GATTCACACCAGTGGGAACACCTTGCTAAGTTTGCATAACGCGCAAATTTGATTTGCTACAA 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968 GATTCACACCAGTGGGAACACCTTGCTAAGTTTGCATAACGCGCAAATTTTATTTGCTACAA 1027

Qy    841 CGCACGCCAGAGGTTGCCCGCAGCGCGCCGCCACCCGTTATTAGATTTAGATTCAAGACACACGCG 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028 CGCACGCCAGAGGTTGCCCGCAGCGCGCCGCCACCCGTTATTAGATTTGATCAAGACACGCG 1087

Qy    901 TTGACGCCCCATCCACCGCCAAAAACAGGCGTATGGGTGTGGACATTACCCACTTCAGTGCTG 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 TTGACGCCCCATCCACCGCCAAAAACAGGCGTATGGTGTGCACATTACCCACTTCAGTGCTG 1147

Qy    961 TTTATCGGCGGACGGATACTAATCTGGCAAATCTCGGGCGGGCACTGGAGCTCAACTGG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148 TTTATCGGCGGACACGGATTACTAATCTGGCAAATCTCGGGCGGGCACTGGAGCTCAACTGG 1207

Qy   1021 ACGCTTCCCGGTCAGCCGGATAACAGCGCGCCACGGTGCTGAACTGGTTGAACGCGCTGG 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208 ACGCTTCCCGGTCAGCCGGATAACAGCGCGCCACGGTGCTGAACTGGTTTGAACGCGCTGG 1267

Qy   1081 CGTCGGCTAAGGGATAACAGCCAGTGGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268 CGTCGGCTAAGCCATAACCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1327

Qy   1141 CAGATGCGTGATAAAAGCCGCGTGTCAGTTAAATACGCCGCCCGGAGGGTGAAAACTGACC 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328 CAGATGCGGTGATAAAACGCCGCTGTCATTAAATACGCCGCCCGGAGGGTGAAAACTGACC 1387

Qy   1201 CTGGCACGGATGTGAAGAGCCGAAATGCGCAGGGCCATGTGTGTTCGTTGGCAGGTTTTACGGCAA 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388 CTGGCAGGGATGTGAAAGAGCGGCGAAATGCGCAGGCATGTTCGTTGGCAGGTTTTACGGCAA 1447

Qy   1261 ATCGTGAATGAAGCACGCATACCGGGCCTGCAGTTTGAGAT 1300
          ||||||||||||||||||||||||||||||||||||||||
Db   1448 ATCGTGAAATGAAGCACGCATACCGGGCCTGCAGTTTGTAAT 1487
```

THIS PAGE LEFT BLANK

```
                                        GenCore version 5.1.9
                       Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on:        June 14, 2006, 10:00:16 ; Search time 889.897 Seconds
                                                  (without alignments)
                                                  10248.037 Million cell updates/sec

Title:
Perfect score:    1308
Sequence:      1 atgaaagcgatccttaatccc.............gcagtttgagatccatcta 1308

Scoring table:    IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0
Searched:      5244920 seqs, 348612423l residues
Total number of hits satisfying chosen parameters:        10489840

Post-processing: Minimum Match 0%
Minimum DB seq length: 0
Maximum DB seq length: 200000000
               Listing first 45 summaries

Database :      N_Geneseq_8:*
               1:   geneseqnl980s:*
               2:   geneseqnl990s:*
               3:   geneseqn200s:*
               4:   geneseqn2001as:*
               5:   geneseqn2001bs:*
               6:   geneseqn2002as:*
               7:   geneseqn2002bs:*
               8:   geneseqn2003as:*
               9:   geneseqn2003bs:*
              10:   geneseqn2003cs:*
              11:   geneseqn2003ds:*
              12:   geneseqn2004as:*
              13:   geneseqn2004as:*
              14:   geneseqn2004bs:*
              15:   geneseqn2005s:*
              15:   geneseqn2006s:*
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1308	100.0	1308	10	ADC87742	Adc87742 DNA encod
2	1308	100.0	1308	12	AAD25463	Aad25463 Escherich
3	1308	100.0	1308	15	ADA19449	Ada19449 E. coli K
4	1296	99.1	1296	14	ADC87744	Adc87744 DNA encod
5	1294.2	98.9	1296	15	AED50817	Aed50817 Escherich
6	1268.6	97.0	1901	15	AEE75422	Aee75422 E. coli B
7	1268	97.0	1901	12	ADO50296	Ado50296 Escherich
8	1267	96.9	1901	15	ADO50295	Ado50295 Escherich
9	1267	96.9	1901	15	AEE75416	Aee75416 E. coli P
10	1265	96.7	1323	5	AAC88885	Aac88885 Escherich
11	1265	96.7	1323	6	AAD36473	Aad36473 Escherich
12	1265	96.7	1323	8	AAD25460	Aad25460 Escherich
13	1265	96.7	1323	12	ADA19445	Ada19445 E. coli B
14	1265	96.7	1323	12	ADO50291	Ado50291 Escherich
15	1265	96.7	1323	12	ADO50301	Ado50301 Kangaroo
16	1265	96.7	1323	15	ADO50303	Ado50303 Escherich
17	1264	96.6	1299	3	ABK12514	Abk12514 DNA encod
18	1264	96.6	1299	8	ACA19297	Aca19297 Prokaryot
19	1264	96.6	1299	12	ADL16137	Adl16137 Escherich
20	1263.8	96.6	1901	6	AAD25463	Aad25463 Escherich
21	1263.8	96.6	1901	8	ADA19449	Ada19449 E. coli K
22	1263.8	96.6	1901	10	ADC87744	Adc87744 DNA encod
23	1263.8	96.6	1901	12	ADO50297	Ado50297 Escherich
24	1263.8	96.6	1901	14	AED50819	Aed50819 Escherich
25	1263.8	96.6	1901	15	AEE75418	Aee75418 E. coli B
26	1261.8	96.6	1901	15	AEE75417	Aee75417 E. coli p
27	1260.8	96.4	1296	12	ADL16139	Adl16139 Shigella
28	1260	96.3	1304	14	AED50824	Aed50824 Escherich
29	1259.2	96.3	3470	3	AAC68298	Aac68298 R15/APPA
30	1259.2	96.3	4060	3	AAC68296	Aac68296 R15/APPA
31	1259.2	96.3	6116	3	AAC68297	Aac68297 R15/APPA
32	1259.2	96.3	6708	3	AAC68295	Aac68295 R15/APPA
33	1259.2	96.3	17732	3	AAC68300	Aac68300 Lama2/APP
34	1259.2	96.3	20623	3	AAC68294	Aac68294 Lama2/APP
35	1258.4	96.2	1315	10	ADL02197	Adl02197 DNA encod
36	1251	95.6	1901	15	AEF15218	Aef15218 Escherich
37	1249.6	95.5	1299	15	ADA19452	Ada19452 E. coli K
38	1243.2	95.3	5421	3	AAC68299	Aac68299 SV40/APPA
39	1243.2	95.0	1296	12	ADL16138	Adl16138 Shigella
40	1243.2	95.0	1299	13	ADW76362	Adw76362 Phytase A
41	1233.6	94.3	1486	4	AAD06831	Aad06831 E. coli a
42	1228	93.9	1489	8	AAA28216	Aaa28216 E. coli a
43	1228	93.9	1489	8	ACC57672	Acc57672 Escherich
44	1222.4	93.5	1486	4	AAD06832	Aad06832 E. coli a
45	1222.4	93.5	1486	8	ACC57673	Acc57673 Escherich

ALIGNMENTS

```
RESULT 1
ADC87742
ADC87742 standard; DNA; 1308 BP.

XX
ID     ADC87742 standard; DNA; 1308 BP.
XX
AC     ADC87742;
XX
DT     01-JAN-2004  (first entry)
XX
DE     DNA encoding the modified Escherichia coli phytase.
XX
KW     Phytase; food supplement; enzyme delivery matrix; soybean meal;
KW     thermotolerance; thermostability; kernel; phytate;
KW     myo-inositol-hexaphosphate; inositol; inorganic phosphate;
KW     thermotolerant; feed value; digestion; gene; ds.
XX
OS     Escherichia coli.
OS     Synthetic.
XX
FH     Key             Location/Qualifiers
FT     CDS             1..1308
FT                     /*tag= a
FT                     /product= "Modified phytase"
XX
PN     US2003103958-A1.
XX
PD     05-JUN-2003.
XX
PF     24-MAY-2002; 2002US-00156660.
XX
PR     13-AUG-1997; 97US-00910798.
PR     01-MAR-1999; 99US-00259214.
PR     13-APR-1999; 99US-00291931.
PR     25-MAY-1999; 99US-00318528.
PR     25-MAY-2000; 2000US-00580515.
PR     24-MAY-2001; 2001US-00866379.
XX
PA     (DIVE-) DIVERSA CORP.
XX
PI     Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI     Mathur EJ;
XX
```

```
DR   WPI; 2003-787039/74.
DR   P-PSDB; ADC87743.
XX
PT   New nucleic acid encoding a polypeptide having phytase activity, useful
PT   in improving the feeding value of phytate rich ingredients or as an aid
PT   in phytate digestion.
XX
PS   Claim 9; SEQ ID NO 1; 113pp; English.
XX
CC   The invention discloses a new isolated or recombinant nucleic acid which
CC   encodes a polypeptide having a phytase activity. Also claimed is a
CC   nucleic acid probe, an amplification primer sequence pair, an expression
CC   cassette comprising the nucleic acid, a vector comprising the nucleic
CC   acid, a transgenic non-human animal or plant, or its seed, comprising the
CC   nucleic acid, an antisense oligonucleotide, inhibiting the translation of
CC   a phytase message in a cell, a heterodimer comprising the polypeptide and
CC   a second domain, an array comprising immobilised polypeptide or nucleic
CC   acid, a hybridoma comprising an antibody that specifically binds to the
CC   polypeptide, a food supplement for an animal, an edible carrier and the
CC   matrix, an edible pellet comprising a granule edible carrier and the
CC   polypeptide, a feed composition, a soybean meal, an edible enzyme delivery
CC   polypeptide, determining whether a compound binds to the polypeptide,
CC   identifying a modulator; whole cell engineering of new or modified
CC   phenotypes by using real-time metabolic flux analysis, increasing
CC   thermotolerance or thermostability of the phytase polypeptide, increasing
CC   the resistance of the phytase polypeptide to enzymatic inactivation in a
CC   digestive system and processing of corn and sorghum kernels. The phytase
CC   activity comprises catalysis of phytate (myo-inositol-hexaphosphate) to
CC   inositol and inorganic phosphate and is thermotolerant. The nucleic acid
CC   is useful in improving the feeding value of phytate rich ingredients or
CC   as an aid in phytate digestion. The sequence presented is the DNA
CC   encoding the modified Escherichia coli phytase.
XX
SQ   Sequence 1308 BP; 315 A; 345 C; 357 G; 291 T; 0 U; 0 Other;

Query Match             100.0%;  Score 1308;  DB 10;  Length 1308;
Best Local Similarity   100.0%;
Matches  1308;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTCTTCTGATTCGGTTAACCCCCCCAATCTGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTCTTCTGATTCGGTTAACCCCCCCAATCTGCA 60

Qy         61 TTCGGCTTCAGAGGTGAGCCGGAGCTGAAAGTGTGGTGTCATTGTCAGTTCGTCATNGGT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGGCTTCAGAGGTGAGCCGGAGCTGAAAGTGTGGTGTCATTGTCAGTTCGTCATNGGT 120

Qy        121 GTGGCTGCGTCCAACCAAGGCCAACGCAACTGATGCAGATGTCACCCCGACGACGCATGGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTGGCTGCTCCAACCAAGGCCACGCAACTGATGATGCACGAAAATGGCCATCCCAGACCA 180

Qy        181 ACCTCGCCCGGTAAAAACTGGGGTGAGCTGACACCGGCCGCGGGGTGCAGTAATCGCCTATCTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCTCGCCCGGTAAAAACTGGGGTGAGCTGACACCGGCCGCGGGGTGCAGTAATCGCCTATCTC 240

Qy        241 GGACATTACTNGGCGTCAGGTGGTAGCCGACGGATTGCTGCCCTAAATGTGGGCTGCCCG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGACATTACTNGGCGTCAGGTGGTAGCCGACGGATTGCTGCCCTAAATGTGGGCTGCCCG 300

Qy        301 CAGTCTCGGTCAGGTCCGGATTATTTGCTGATGTCGACGAGCGCTACCCGTAAAACAGGCGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGTCTGGTCAGGTCCGGATTATTGCTGATGTCGACGAGCGCTACCCGTAAAACAGGCGAA 360

Qy        361 GCCTTCGCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTTCGCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420

Qy        421 TCCAGTCCCGATCCGGTTATTTAATCCTCTAAAAACTGGGGTGGCGTTGCCAACTGGATAACGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCAGTCCCGATCCGGTTATTTAATCCTCTAAAAACTGGGGTGGCGTTGCCAACTGGATAACGCG 480

Qy        481 AACGTGACTGACGCGGATCCTCGAGAGGGGCACGAGGGGTCAATTGCTGACTTTACCGGGGCAT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACGTGACTGACGCGGATCCTCGAGAGGGGCACGAGGGGTCAATTGCTGACTTTACCGGGGCAT 540

Qy        541 TATCAAACGGGCGGTTTCCGGACACTGGACAGGGGTGCCTTAATTTTCCGGCAATCAAAACTGTGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TATCAAACGGGCGGTTTCCGGACACTGGACAGGGGTGCCTTAATTTTCCGGCAATCAAAACTGTGC 600

Qy        601 CTTAAACGTGAGAAAACAGGACGAAAAGCTTGTTCATTACGCAGGCGATTACCATCGGAACTC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTAAACGTGAGAAAACAGGACGAAAAGCTTGTTCATTACGCAGGCGATTACCATCGGAACTC 660

Qy        721 GAGATATTTCTCCCTGCAACAAGCACAGGGAATGCCCGGAGCCCGGGGTGGGGAAGGGATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGATATTTCTCCCTGCAACAAGCACAGGGAATGCCCGGAGCCCGGGGTGGGGAAGGGATCACC 780

Qy        661 AAGGTGAGGCGCCGACTGTGTCTCATTAAACCGGTGCGTGAAGGCCTCGGACATCAATGCTGACG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGGTGAGGCGCCGACTGTGTCTCATTAAACCGGTGCGTGAAGGCCTCGGACATCAATGCTGACG 720

Qy        781 GATTCACCACCAGTGGAACACCTTGCTAAGTTTGCATAAGCGGCCAAATTTGCATTGCTACAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GATTCACCACCAGTGGAACACCTTGCTAAGTTTGCATAACGGGCCAAATTTGCATTTGCTACAA 840

Qy        841 CGCACGCCAGAGGTTGCCCGGAGCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CGCACGCCAGAGGTTGCCCGGAGCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 900

Qy        901 TTGACGCCCCATCCACCGCAAAAACAGGCGTPATGGGTGGACATTACCCACTTCAGTCGTCTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTGACGCCCCATCCACCGCAAAAACAGGCGTAAGGGTGGACATTACCCACTTCAGTGGTCTG 960

Qy        961 TTTATCGCCGGACACGATACTRAATCTGGCAAATTCTCGGCGGACAGTCAAGTCGCTCAACTGG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTTATCGCCGGACACGATACTRAATCTGGCAAATTCTCGGCGGACAGTCAAGTCGCTCAACTGG 1020

Qy       1021 ACGCTTCCCGGTCAGCCGGATAACACGCGCCCAGGTGGTGGTTGGGAACTGGTGTTGGTGAACGCGTGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACGCTTCCCGGTCAGCCGGATAACACGCGCCCAGGTGGTGGTTGGGAACTGGTGTTGGTGAACGCGTGG 1080

Qy       1081 CGTCGGCTAACGGATAACAGCCAGTGGATTCAGGTTTCGCTGCTGGTCTTCCCAGACTTTACAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CGTCGGCTAACGGATAACAGCCAGTGGATTCAGGTTTCGCTGCTGGTCTTCCCAGACTTTACAG 1140

Qy       1141 CAGATGGGTGATAAAAACGCCCGCTGTCATTAAATACCGGTCAGAGAGGTGAAAACTGACC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CAGATGGGTGATAAAAACGCCCGCTGTCATTAAATACCGGTCAGAGAGGTGAAAACTGACC 1200

Qy       1201 CTGGCAGGATGTGAAGAGGCGAAATGCCGCAGGCCGCATGGTGCAGTGTTCGGCAGGGTTTTTACGGCAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CTGGCAGGATGTGAAGAGGCGAAATGCCGCAGGCCGCATGGTGCAGTGTTCGGCAGGGTTTTTACGGCAA 1260

Qy       1261 ATCGTGAATGAATGAAGCAGCGCATACCGGGCGTGCAGTTTGAGATCTCATCTA 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ATCGTGAATGAATGAAGCAGCGCATACCGGGCGTGCAGTTTGAGATCTCATCTA 1308
```

```
RESULT 2
AD050299
ID   AD050299 standard; DNA; 1308 BP.
XX
AC   AD050299;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Escherichia coli K12 appA phytase mutant 819PH59 DNA.
XX
KW   AppA phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW   phytate; animal feed; fish feed; dough; baking; gene; mutant; ds.
XX
OS   Escherichia coli; K12.
XX
OS   Synthetic.
XX
FH   Key                 Location/Qualifiers
```

```
FT   CDS              1..1296
FT                    /product= "Mutant phytase"
FT                    /note= "Does not include stop codon"
FT                    /partial
FT   mutation
FT                    /replace(202,T)
FT                    /*tag= a
FT                    sequence"
FT   mutation
FT                    /replace(203,G)
FT                    /*tag= b
FT                    /note= "Corresponds to position 389 in the wild type
FT                    sequence"
FT   mutation
FT                    /replace(250,C)
FT                    /*tag= c
FT                    /note= "Corresponds to position 390 in the wild type
FT                    sequence"
FT   mutation
FT                    /replace(251,A)
FT                    /*tag= d
FT                    /note= "Corresponds to position 437 in the wild type
FT                    sequence"
FT   mutation
FT                    /replace(252,A)
FT                    /*tag= e
FT                    /note= "Corresponds to position 438 in the wild type
FT                    sequence"
FT   mutation
FT                    /replace(255,C)
FT                    /*tag= f
FT                    /note= "Corresponds to position 439 in the wild type
FT                    sequence"
FT   mutation
FT                    /replace(283,G)
FT                    /*tag= g
FT                    /note= "Corresponds to position 442 in the wild type
FT                    sequence"
FT   mutation
FT                    /replace(285,G)
FT                    /*tag= h
FT                    /note= "Corresponds to position 470 in the wild type
FT                    sequence"
FT   mutation
FT                    /replace(289,G)
FT                    /*tag= i
FT                    /note= "Corresponds to position 472 in the wild type
FT                    sequence"
FT   mutation
FT                    /replace(290,A)
FT                    /*tag= j
FT                    /note= "Corresponds to position 476 in the wild type
FT                    sequence"
FT   mutation
FT                    /replace(291,G)
FT                    /*tag= k
FT                    /note= "Corresponds to position 477 in the wild type
FT                    sequence"
FT   mutation
FT                    /replace(502,A)
FT                    /*tag= l
FT                    /note= "Corresponds to position 478 in the wild type
FT                    sequence"
FT   mutation
FT                    /replace(503,G)
FT                    /*tag= m
FT                    /note= "Corresponds to position 689 in the wild type
FT                    sequence"
FT   mutation
FT                    /replace(504,C)
FT                    /*tag= n
FT                    /note= "Corresponds to position 690 in the wild type
FT                    sequence"
FT   mutation
FT                    /replace(541,C)
FT                    /*tag= o
FT                    /note= "Corresponds to position 691 in the wild type
FT                    sequence"
FT   mutation
FT                    /replace(542,G)
FT                    /*tag= p
FT                    /note= "Corresponds to position 728 in the wild type
FT                    sequence"
FT   mutation
FT                    /replace(543,G)
FT                    /*tag= q
FT                    /note= "Corresponds to position 729 in the wild type
FT                    sequence"
FT   mutation
FT                    /replace(544,G)
FT                    /*tag= r
FT                    /note= "Corresponds to position 730 in the wild type
FT                    sequence"
FT   mutation
FT                    /replace(676,A)
FT                    /*tag= s
FT                    /note= "Corresponds to position 863 in the wild type
FT                    sequence"
FT   mutation
FT                    /replace(677,A)
FT                    /*tag= t
FT                    /note= "Corresponds to position 864 in the wild type
FT                    sequence"
FT   mutation
FT                    /replace(829,T)
FT                    /*tag= u
FT                    /note= "Corresponds to position 1016 in the wild type
FT                    sequence"
XX
PN   US2004091968-A1.
XX
PD   13-MAY-2004.
XX
PF   20-JUN-2003; 2003US-00601319.
XX
PR   13-AUG-1997; 97US-00910798.
PR   01-MAR-1999; 99US-00259214.
PR   13-APR-1999; 99US-00291931.
PR   25-MAY-1999; 99US-00318528.
PR   25-MAY-2000; 2000US-00580515.
PR   24-MAY-2001; 2001US-00866379.
XX
PA   (SHOR/)  SHORT J M.
PA   (KRBT/)  KRETZ K.
PA   (GRAY/)  GRAY K A.
PA   (BART/)  BARTON N R.
PA   (GARR/)  GARRETT J B.
PA   (ODON/)  O'DONOGHUE E.
PA   (MATH/)  MATHER E J.
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI   Mather EJ;
XX
DR   WPI; 2004-374952/35.
DR   P-PSDB; ADO50300.
XX
PT   Producing phytase, involves providing nucleic acid derived from bacteria
PT   encoding polypeptide having phytase activity, and expressing nucleic acid
PT   in yeast.
XX
PS   Claim 3; SEQ ID NO 9; 74pp; English.
XX
CC   The invention relates to a method for producing phytase that involves
CC   providing a nucleic acid encoding phytase derived from a bacteria, and
CC   expressing the nucleic acid in a yeast under conditions that allow
CC   expression of the enzyme in the yeast. The invention also relates to
CC   modified phytase enzyme which has improved thermal tolerance and protease
CC   stability at low pH. The phytase enzyme can be used in foodstuffs to
CC   improve the feeding value of phytate rich ingredients, and in diet of
CC   numerous animals including mammals, fowls and fishes, commercially
CC   significant mammals such as pigs, goats, laboratory rodents, commercially
CC   significant avian species such as chicken, ducks, doves, parrot, etc.,
CC   commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC   etc., in dough making and baking, in dietary aids for animals. The method
CC   provides easy manufacture of the active ingredient loaded biocompatible
CC   composition, higher yields and loading efficiency. The phytase
CC   incorporated in to the dietary aid is safe for animals. The present
CC   sequence is Escherichia coli K12 appA phytase mutant DNA.
XX
SQ   Sequence 1308 BP; 315 A; 345 C; 357 G; 291 T; 0 U; 0 Other;

Query Match             100.0%;  Score 1308;  DB 12;  Length 1308;
Best Local Similarity   100.0%;
Matches  1308;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGAAAGCGGATCTTAATCCACCATTTTTATCTCTTCGTTGATTCCGTTAACCCGCAAATCTGCA 60
              ||||||||||||||||||||  |||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTCGTTGATTCCGTTAACCCGCAAATCTGCA 60
```

```
Qy   61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGTGTGATTGTCAGTCGTCATGGT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTTACTGGCGTCAGCCGTCAGCCGTCAGCGGCGTCTGGGTAG 120

Qy  121 GTGGCGTGCTCCAACCAAGCAGCCACCGCACCGACTGATTGTCAGTCGTCATGGT 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTGGCGTGCTCCAACCAAGCCACGCAACTGATGCGGAGATGTCACCCCAGAGACGGCATGGCCA 180

Qy  181 ACCTGGCCGGTAAAACTGGGTGAGCTGACACCGCCGCCGCCTGGTCAATCGCCTATCTC 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ACCTGGCCGGTAAAACTGGGTGAGCTGACACCGCCGCCTGGTGAGCTAATCGCCTATCTC 240

Qy  241 GGACATTACTGGCGTCAGCCGTCTCGGTGAGCCGACGGGATTGCTGCCTAAATGTGGCTGCCCG 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GGACATTACTGGCGTCAGCCGTCTCGGTGAGCCGACGGGATTGCTGCCTAAATGTGGCTGCCCG 300

Qy  301 CAGTCTGGTCAGGTCGGCGATTATTTGCTGATGTCGACGAGCGTACCCGGTAAAAACAGGCGAA 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGTCTGGTCAGGTCGGCGATTATTTGCTGATGTCGACGAGCGTACCCGGTAAAAACAGGCGAA 360

Qy  361 GCCTTCGCCGCCGGCCTGGCCACCTGACGTGTGCAATAACCGTACATACCCAGGCAGATACG 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GCCTTCGCCGCCGGCCTGGCCACCTGACGTGTGCAATAACCGTACATACCCAGGCAGATACG 420

Qy  421 TCCAGTCCCGGATCCCGGTTATTTAATCCTCTAAAAACTGGGCGTTTGCCAACTGCAGGCGA 480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TCCAGTCCCGGATCCCGGTTTATTTAATTCTCTAAAAACTGGGCGTTTGCCAACTGGAATAACGCG 480

Qy  481 AACGTGACTGACGCGATCCTCGAGGACGGGGCAGGAGGGGTCAATTGCTGACTTTTACCGGGCAT 540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 AACGTGACTGACGCGATCCTCGAGGACGGGGCAGGAGGGGTCAATTGCTGACTTTTACCGGGCAT 540

Qy  541 TATCAAAACGGCGTTTCGCGAACTGGAACGCTTAATTTTTCCGCAAATCAAAACTTGTGC 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 TATCAAAACGGCGTTTCGCGAACTGGAACGCTTAATTTTTCCGCAAATCAAAACTTGTGC 600

Qy  601 CTTAAACGTGAGAAACAGGACGGAAAAGCTGTTCATTAACGCACATTACCATCGGAACTC 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CTTAAACGTGAGAAACAGGACGGAAAAGCTGTTCATTAACGCACATTACCATCGGAACTC 660

Qy  661 AAGGTGAGCGCCGACTGTGTTCTCATTAAACCGGTAAGCCTCCGCGTAAGTCTGACG 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 AAGGTGAGCGCCGACTGTGTTCTCATTAAACCGGTAAGCCTCCGCGTAAGTCTGACG 720

Qy  721 GAGATATTTCTCCTGCAACAAGACACAGGGATGCCGGAGCCGCGGGAGGAGGATCACC 780
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GAGATATTTCTCCTGCAACAAGACACAGGGATGCCGGAGCCGCGGGAGGAGGATCACC 780

Qy  781 GATTCACACCAGTGGAACACCTTGCTAAGTTGCATAACGCGCAAATTTGATTTGCTACAA 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GATTCACACCAGTGGAACACTAAATCTGGCAAATCTCGGCGCACTGCCCCACTTACCGCTG 840

Qy  841 CGCACGCCCAGAGGTTGCCCGCGCCAGCCCCGCCGTTATTTGATTGATCAAGACAGCG 900
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 CGCACGCCCAGAGGTTGCCCGCGCCAGCCCCGCCGTTATTTGATTTGATCAAGAGATCACG 900

Qy  901 TTGACGCCCCATCCACGCGCAAAAACAGGCGGTATGGTGTGACATTACCCACTTCAGTCGCTG 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TTGACGCCCCATCCACGCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTCGCTG 960

Qy  961 TTTATCGCCGGACACGATACTAAATCTGGCAAAATCTCGGCGGACGCACTGCCCGGAGCTCAACTGG 1020
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 TTTATCGCCGGACACGATACTAAATCTGGCAAATCTCGGCGGACGCACTGCCCGGAGCTCAACTGG 1020

Qy 1021 ACGCTTCCCGGTCAGCGGACTAAACACGCCGGGTGGTGAACACGCGACGCCTGG 1080
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 ACGCTTCCCGGTCAGCGGATAACACGCCGGGTGGTGAACACGCGACGCCTGG 1080

Qy 1081 CGTCCGGCTAAGCGATAACAGCCAGGTTTCGCTCGGTCTTCTTCCAGACTTTACAG 1140
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 CGTCCGGCTAAGCGATAACAGCCAGGTTTCGCTCGGTCTTCTTCCAGACTTTACAG 1140

Qy 1141 CAGATCGTGATAAAACGCCCGCTGTCATTAAATACCCGCCCGGAGGAGGTGAAACTGACC 1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 CAGATCGTGATAAAACGCCCGCTGTCATTAAATACCCGCCCGGAGGAGGTGAAACTGACC 1200
```

```
Qy 1201 CTGGCAGGATGTGAAGAGCGGAAATGCCGCAGGCGCATGTTGTTCGTTGGCAGGTTTTTACGGCAA 1260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 CTGGCAGGATGTGAAGAGCGGAAATGCCGCAGGCGCATGTTGTTCGTTGGCAGGTTTTTACGGCAA 1260

Qy 1261 ATCGTGAATGAAGCACCCATACCGGCGGCCTGCCAGTTTGAGATCTCATCTA 1308
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 ATCGTGAATGAAGCACCCATACCGGCGGCCTGCCAGTTTGAGATCTCATCTA 1308
```

RESULT 3
ID   AEE75420 standard; DNA; 1308 BP.
XX
AC   AEE75420;
XX
DT   23-FEB-2006 (first entry)
XX
DE   E. coli B modified appA phytase 819PH59 DNA sequence SEQ ID NO: 9.
XX
KW   ds; coding sequence; mutant; thermostable; phytase;
KW   protein stabilization; pharmaceutical; fertilizer; cramp; cell culture;
KW   osteoporosis; osteopathic; food; appA.
XX
OS   Synthetic.
XX
OS   Escherichia coli K12.
XX
FH   Key   Location/Qualifiers
FH
FT   CDS   1..1296
FT   /*tag= a
FT   /product= "modified phytase appA"
XX
PN   US2005281792-A1.
XX
PD   22-DEC-2005.
XX
PF   01-SEP-2004; 2004US-00933115.
XX
PR   13-AUG-1997; 97US-00910798.
PR   01-MAR-1999; 99US-00259214.
PR   13-APR-1999; 99US-00291931.
PR   25-MAY-1999; 99US-00318528.
PR   25-MAY-2000; 2000US-00580515.
PR   24-MAY-2001; 2001US-00866379.
XX
PA   (SHOR/) SHORT J M.
PA   (KRET/) KRETZ K A.
PA   (GRAY/) GRAY K A.
PA   (BART/) BARTON N R.
PA   (GARR/) GARRETT J B.
PA   (ODON/) ODONOGHUE E.
PA   (BAUM/) BAUM W.
PA   (ROBE/) ROBERTSON D E.
PA   (ZORN/) ZORNER P.
XX
PI   Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, Odonoghue E;
PI   Baum W, Robertson DE, Zorner P;
XX
DR   WPI; 2006-055022/06.
DR   P-PSDB; AEE75421.
XX
PT   Formulation useful as dietary supplement for treating, preventing or
PT   reversing osteoporosis and bone loss, and preventing muscle cramps,
PT   comprises one or more polypeptides having phytase activity.
XX
PS   Example 2; SEQ ID NO 9; 82pp; English.
XX
CC   The present sequence is that of a polynucleotide encoding a novel
CC   modified phytase derived from the Escherichia coli appA protein, with
CC   improved thermal tolerance and protease stability compared to the wild-
CC   type. The present invention relates to a novel formulation, comprising at

```
CC   least one polypeptide having phytase activity, where the polypeptide is a
CC   modified sequence (AEE75421) derived from the Escherichia coli K-12 appA
CC   phytase (AEE75419). The modification of the enzyme, comprising a series
CC   of mutations, improves the thermal tolerance and protease stability of
CC   the protein. The specification also claims a pharmaceutical composition;
CC   a kit; an immobilized phytase; a fertilizer or soil additive; a liquid
CC   supplement for preventing muscle cramps; a hydrating agent; a tissue
CC   culture or cell culture media; and a plant food additive, all comprising
CC   the novel phytase, and a method of reducing pollution and increasing
CC   nutrient availability in an environment or environmental sample by
CC   degrading environmental phytic acid. The novel enzyme is a dietary
CC   supplement useful for treating, preventing or reversing osteoporosis or
CC   bone loss, and preventing muscle cramps. The liquid supplement is useful
CC   for preventing muscle cramps. The method is useful for reducing pollution
CC   and increasing nutrient availability in an environment or environmental
CC   sample by degrading environmental phytic acid, where the environment or
CC   environmental sample comprises a soil or a body of water. The immobilized
CC   phytase is useful in foodstuffs for improving the feeding value of
CC   phytate rich ingredients.
XX
SQ   Sequence 1308 BP; 315 A; 345 C; 357 G; 291 T; 0 U; 0 Other;

Query Match             100.0%;  Score 1308;  DB 15;  Length 1308;
Best Local Similarity   100.0%;  Pred. No. 0;  Mismatches 0;
Matches 1308;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 ATGAAAACGGATCTTAATCCATTTTACTCTTCTTCTGATTCCGTTAACCCCGCAAATCTGCA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAACGGATCTTAATCCATTTTACTCTTCTTCTGATTCCGTTAACCCCGCAAATCTGCA 60

Qy     61 TTCGCTCAGAGTGAGCCGGAGCTGAAGGCTGGAAAGTGTGGGAAAGTGTCAGTCGTCATTGGT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCGCTCAGAGTGAGCCGGAGCTGAAGGCTGGAAAGTGTGGGAAAGTGTCAGTCGTCATTGGT 120

Qy    121 GTGCGTGCTCCAACCAAGGCCACGGCAACTGATGCAGGATGTCACCCAGACGGCACATGGCCA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGCGTGCTCCAACCAAGGCCACGGCAACTGATGCAGGATGTCACCCAGACGGCACATGGCCA 180

Qy    181 ACCTGGCCGGGTAAAACTGGGTGAGCTGAGCTAATCCGCCTATCTC 240
          |||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACCTGGCCGGGTAAAACTGGGTGAGCTGAGCTAATCCGCCTATCTC 240

Qy    241 GGACATTACTCGGCGTCAGCGTCGCGTCGGTAGCCGACGGAATGG 300
          |||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGACATTACTCGGCGTCAGCGTCGCGTCGGTAGCCGACGGAATGG 300

Qy    301 CAGTCTGGTCAGGTCGCGGATTAATTGCTGAATGTCGACACGAGCGTACCCGTAAAACAGGGCGAA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGTCTGGTCAGGTCGCGGATTAATTGCTGAATGTCGACACGAGCGTACCCGTAAAACAGGGCGAA 360

Qy    361 GCCTTCGCCGCCGCCGGCTGGCACCTGACTGTGCAATAACCGTACATACCGAGGCCAGAGATACG 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCTTCGCCGCCGCCGGCTGGCACCTGACTGTGCAATAACCGTACATACCGAGGCCAGAGATACG 420

Qy    421 TCCAGTCCCGATCCGGTTATTAATCCTCTAAAAACTGGCCGTTTGCCAACTGGATAACGCG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCCAGTCCCGATCCGGTTATTAATCCTCTAAAAACTGGCCGTTTGCCAACTGGATAACGCG 480

Qy    481 AACGTGACTGACGCGGATCCTCGAGAGGGCGCAGGAGGGTTCAAGTTCATTAGCTTTACCGGGCAT 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AACGTGACTGACGCGGATCCTCGAGAGGGCGCAGGAGGGTTCAAGTTCATTAGCTTTACCGGGCAT 540

Qy    541 TATCAAACGGCCGTTTCGCGACTGGAACGGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TATCAAACGGCCGTTTCGCGACTGGAACGGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600

Qy    601 CTTAAAACGTGAGAAACAGGAGCGAAAAGCTGTTCATTAACGCAGGCAATTACCATCGGAACTC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTTAAAACGTGAGAAACAGGAGCGAAAAGCTGTTCATTAACGCAGGCAATTACCATCGGAACTC 660

Qy    661 AAGGTGAGCGCCGACTGTGTCTCATTAACCGGTGCGGTAAGCCTCGCACTGTATTAACGG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AAGGTGAGCGCCGACTGTGTCTCATTAACCGGTGCGGTAAGCCTCGCACTGTATTAACGG 720

Qy    721 GAGATATTTCTCCTGCCAACAAGCACAGGAGAATGCCGGACCCGGGGTCGGGGGAAGGATCACC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GAGATATTTCTCCTGCCAACAAGCACAGGAGAATGCCGGACCCGGGGTCGGGGGAAGGATCACC 780

Qy    781 GATTCACACCAGTGGAAACACCTTGCTAAGTTTGCATAAGACGGCGCAATTTGATTTGCTACAA 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GATTCACACCAGTGGAAACACCTTGCTAAGTTTGCATAAGACGGCGCAATTTGATTTGCTACAA 840

Qy    841 CGCACGCCAGAGGGTTGCCCGCGCCAGCCCCGCCCCGTTATTAGATTTGATCAAGACAGCG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CGCACGCCAGAGGGTTGCCCGCGCCAGCCCCGCCCCGTTATTAGATTTGATCAAGACAGCG 900

Qy    901 TTGACGGCCCCATCCACCGGCAAAAAACAGGGGTATGGTGTGACATTACCCACTTTCAGTGCTG 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTGACGGCCCCATCCACCGGCAAAAAACAGGGGTATGGTGTGACATTACCCACTTTCAGTGCTG 960

Qy    961 TTTATCGGCCGGGACACGATACTAATCTGGCAAAATCTCGGCGGGCCGACTGGAGCTCAAACTGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTTATCGGCCGGGACACGATACTAATCTGGCAAAATCTCGGCGGGCCGACTGGAGCTCAAACTGG 1020

Qy   1021 ACGCTTCCCGGGTCAGCCGGATAACACGCGCCGCCCAGGTGGTGAAACTGGGTGTTTGAACGGCTGG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGCTTCCCGGGTCAGCCGGATAACACGCGCCGCCCAGGTGGTGAAACTGGGTGTTTGAACGGCTGG 1080

Qy   1081 CGTCGGCGTAAGCGGATAACAGCCCAGTGGGATTCAGGTTTCCGCTGGTTTCCCAGACTTTACAG 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CGTCGGCGTAAGCGGATAACAGCCCAGTGGGATTCAGGTTTCCGCTGGTTTCCCAGACTTTACAG 1140

Qy   1141 CAGATGCGTGATAAAAACGCGCGTCATTAAAAATACGCGCCCCCGAGAGGTGAAACTGACC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAGATGCGTGATAAAAACGCGCGTCATTAAAAATACGCGCCCCCGAGAGGTGAAACTGACC 1200

Qy   1201 CTGGCAGGATGTGAAGAGGCGAAATGCGCAGGGCATGTGTTCGTTGGTTTACGGCAA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGCAGGATGTGAAGAGGCGAAATGCGCAGGGCATGTGTTCGTTGGTTTACGGCAA 1260

Qy   1261 ATCGTGAATGAAGCACACGATACCCGGCGTGCGAGTTTGACATCTCATCTA 1308
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAGCACACGATACCCGGCGTGCGAGTTTGACATCTCATCTA 1308

RESULT 4
AED50817
ID   AED50817 standard; DNA, 1296 BP.
XX
AC   AED50817;
XX
DT   29-DEC-2005  (first entry)
XX
DE   Escherichia coli phytase appA mutant DNA SEQ ID NO: 1 #1.
XX
KW   Transgenic animal; gene therapy; feedstuff; food; thermostable; mutant;
KW   ds; coding sequence; 3-phytase; recombinant DNA; appA gene.
XX
OS   Escherichia coli; strain K12.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FH
FT   CDS             1..1296
FT                   /*tag= a
FT                   /EC_number= "3.1.3.8"
FT                   /partial
FT                   /product= "Escherichia coli phytase mutant protein"
FT                   /note= "No stop codon"
XX
PN   US2005246780-A1.
XX
PD   03-NOV-2005.
XX
PF   11-FEB-2005; 2005US-0005354.
XX
PR   13-AUG-1997; 97US-00910798.
```

```
PR   01-MAR-1999;
PR   13-APR-1999;   99US-00259214.
PR   25-MAY-1999;   99US-00291931.
PR   25-MAY-2000;   99US-00318528.
PR   24-MAY-2001;   2000US-00580515.
PR   24-MAY-2001;   2001US-00866379.
PR   24-MAY-2002;   2002US-00156660.
XX
PA   (SHOR/)  SHORT J M.
PA   (KRET/)  KRETZ K.
PA   (GRAY/)  GRAY K A.
PA   (BART/)  BARTON N R.
PA   (GARR/)  GARRETT J B.
PA   (ODON/)  ODONOGHUE E.
PA   (MATH/)  MATHUR E J.
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, Odonoghue E;
PI   Mathur EJ;
XX
DR   WPI; 2005-733851/75.
DR   P-PSDB; AED50818.
XX
PS   Claim 1; SEQ ID NO 1; 104pp; English.
XX
PT   New isolated or recombinant nucleic acid encoding phytase enzymes, useful
PT   as foodstuff, and for oil degumming, producing an animal feed, delivering
PT   a phytase enzyme supplement to an animal.
XX
CC   The present invention relates to a recombinant phytase (EC 3.1.2.8)
CC   protein and its encoding also nucleic acid molecule. Phytases of the
CC   invention is a thermostable protein. They are useful in foodstuffs, for
CC   oil degumming, producing an animal feed, delivering a phytase enzyme
CC   supplement to an animal, increasing the resistance of a phytase
CC   polypeptide to enzymatic inactivation in a digestive system of an animal
CC   and in processing of corn and sorghum kernels. The invention is also
CC   useful in gene therapy and in production of transgenic animals. The
CC   present sequence is a Escherichia coli phytase appA mutant DNA. Note: The
CC   present sequence is the SEQ ID NO: 1 which is given in the sequence
CC   listing. This sequence differs from the SEQ ID NO: 1 shown on page 19 in
CC   disclosure of the specification (see AED50824).
XX
SQ   Sequence 1296 BP; 311 A; 342 C; 356 G; 287 T; 0 U; 0 Other;

     Query Match             99.1%;  Score 1296;  DB 14;  Length 1296;
     Best Local Similarity   100.0%;  Pred. No. 0;
     Matches 1296;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 ATGAAAAGCGATCTTAATCCCATTTTTATCTCTTCTGAATTCCGTTAACCCCGCCAAATCTGCA     60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAAAGCGATCTTAATCCCATTTTTATCTCTTCTGAATTCCGTTAACCCCGCCAAATCTGCA     60

Qy         61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTGTCAGTCGTCAGGT    120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTGTCAGTCGTCAGGT    120

Qy        121 GTGCCGTGCTCCAACCAAGGCCACGCAACTGATGCGATGTGCACCCCGAGCACGCATGGCCA    180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTGCCGTGCTCCAACCAAGGCCACGCAACTGATGCGATGTGCACCCCGAGCACGCATGGCCA    180

Qy        181 ACCTCGGCCCGGTAAAAACTGGGGTGAGCTGACAGCGATGATGCCTATCTC    240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCTCGGCCCGGTAAAAACTGGGGTGAGCTGACAGCGATGATGCCTATCTC    240

Qy        241 GGACAATTACTGGCGTCAGCGTCGGTAGCCGACGGATTGCTGCCTAAAGTGGGCTGCCCG    300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGACAATTACTGGCGTCAGCGTCGGTAGCCGACGGATTGCTGCCTAAAGTGGGCTGCCCG    300

Qy        301 CAGTCTGGTCAGGTCGCGATTATTGTGATGTCGACGAGCGTACCCGTAAAAACAGGCGAA    360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGTCTGGTCAGGTCGCGATTATTGTGATGTCGACGAGCGTACCCGTAAAAACAGGCGAA    360

Qy        361 GCCTTCGCGCCGCGGCGGCTGGCACCTGACTGTGTCCAATAACCGTACATAACCCAGGCAGATACG    420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTTCGCGCCGCGGCGGCTGGCACCTGACTGTGTCCAATAACCGTACATAACCCAGGCAGATACG    420
```

```
Qy        421 TCCAGTCCCGATCCGGTTATTTAATCCTTCTAAAAAACTGGATAACGCG    480
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCAGTCCCGATCCGGTTATTTAATCCTTCTAAAAAACTGGATAACGCG    480

Qy        481 AACGTGACTGACGCGGATCCTCGGAGAGGGGCAGGAGGGGTCAATTGCTGACTTTACCGGGCAT    540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACGTGACTGACGCGGATCCTCGGAGAGGGGCAGGAGGGGTCAATTGCTGACTTTACCGGGCAT    540

Qy        541 TATCAAACGGGCGTTTCGCGGACTGGAACGGGTGCTTAAATTTTCCGCAAATCAAAACTTGTGC    600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TATCAAACGGGCGTTTCGCGGACTGGAACGGGTGCTTAAATTTTCCGCAAATCAAAACTTGTGC    600

Qy        601 CTTAAACGTGGAGAAAACAGGAGGAAAAGCTGTTCATTAACGCAGGCAATTACCATCGGAACTC    660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTAAACGTGGAGAAAACAGGAGGAAAAGCTGTTCATTAACGCAGGCAATTACCATCGGAACTC    660

Qy        661 AAGGTGAGCGGCCGGACTGTGTCTCATTAAACCGGTGGGGTGGCTTAAATT    720
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGGTGAGCGGCCGGACTGTGTCTCATTAAACCGGTGGGGTGGCTTAAATT    720

Qy        721 GAGATATTTCTCCCTGCAACAAGCACAGGGAATGCCCGGACGCCGGGGTGGGGAGGGATCAA    780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGATATTTCTCCCTGCAACAAGCACAGGGAATGCCCGGACGCCGGGGTGGGGAGGGATCAA    780

Qy        781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCAAATTTGATTTGCTACAA    840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCAAATTTGATTTGCTACAA    840

Qy        841 CGCACGCCAGAGGTTGCCCGGCAGCGCCCACCCCGTTATTAGATTGATCAAGACAGCG    900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CGCACGCCAGAGGTTGCCCGGCAGCGCCCACCCCGTTATTAGATTGATCAAGACAGCG    900

Qy        901 TTGACCCCCATCCACCGCAAAAAACAGGGCGTATGGGTTGGACATTACCCACTTCAGTGCTG    960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTGACCCCCATCCACCGCAAAAAACAGGGCGTATGGGTTGGACATTACCCACTTCAGTGCTG    960

Qy        961 TTTATCGCCGGACACCGATACTAATCTCGGCAAATCTCGGCGCGCCACTGGACAGTGAGTTCGG    1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTTATCGCCGGACACCGATACTAATCTCGGCAAATCTCGGCGCGCCACTGGACAGTGAGTTCGG    1020

Qy       1021 ACGCTTCCCGGTCAGCGCGGATAAACACGCCCGCCAGGTGGTGGTTTGAACGTGG    1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACGCTTCCCGGTCAGCGCGGATAAACACGCCCGCCAGGTGGTGGTTTGAACGTGG    1080

Qy       1081 CGTCGGCTAAACGATAAACAGCCAGTGGATTCAGGTTTCGCTGCGTCTTCCAGACTTTACAG    1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CGTCGGCTAAACGATAAACAGCCAGTGGATTCAGGTTTCGCTGCGTCTTCCAGACTTTACAG    1140

Qy       1141 CAGATGGCGTGATAAAAACGCCCGTCATTAAATACCGCGGAGTTGGCTGGGTTTTACGCCA    1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CAGATGGCGTGATAAAAACGCCCGTCATTAAATACCGCGGAGTTGGCTGGGTTTTACGCCA    1200

Qy       1201 CTGGCAGGATGTGAAGAGGCGAAATGGCCAGGGGCCATGTGTTGGTTCGTTGGCAGGTTTTACGGCAA    1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CTGGCAGGATGTGAAGAGGCGAAATGGCCAGGGGCCATGTGTTGGTTCGTTGGCAGGTTTTACGGCAA    1260

Qy       1261 ATCGTGAATGAAGCACGCCATACCGGCGTGCAGTTTG    1296
              |||||||||||||||||||||||||||||||||||||
Db       1261 ATCGTGAATGAAGCACGCCATACCGGCGTGCAGTTTG    1296
```

```
RESULT 5
AEE75422
ID   AEE75422 standard; DNA; 1901 BP.
XX
AC   AEE75422;
XX
DT   23-FEB-2006  (first entry)
XX
DE   E. coli B modified appA phytase 819PH59 DNA sequence SEQ ID NO: 9.
XX
KW   ds; gene; mutant; thermostable; phytase; protein stabilization;
```

```
KW   pharmaceutical; fertilizer; cramp; cell culture; osteoporosis;
KW   osteopathic; food; appA.
XX
OS   Escherichia coli K12.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FH
FT   CDS             188..1486
FT                   /product= "modified AppA phytase"
FT   mutation        replace(389,T)
FT                   /*tag= a
FT                   /note= "Wild type base T replaced by G"
FT   mutation        replace(390,G)
FT                   /*tag= b
FT                   /note= "Wild type base C replaced by G"
FT   mutation        replace(401..403
FT                   /*tag= d
FT                   /note= "Wild type base G replaced by A"
FT   unsure          401..403
FT                   /note= "ecodes Arg"
FT   mutation        replace(437,C)
FT                   /*tag= e
FT                   /note= "Wild type base C replaced by T"
FT   mutation        replace(438,A)
FT                   /*tag= f
FT                   /note= "Wild type base A replaced by G"
FT   mutation        replace(439,A)
FT                   /*tag= g
FT                   /note= "Wild type base A replaced by G"
FT   mutation        replace(470,G)
FT                   /*tag= h
FT                   /note= "Wild type base A replaced by G"
FT   mutation        replace(472,G)
FT                   /*tag= i
FT                   /note= "Wild type base G replaced by C"
FT   mutation        replace(476,A)
FT                   /*tag= j
FT                   /note= "Wild type base G replaced by T"
FT   mutation        replace(477,A)
FT                   /*tag= k
FT                   /note= "Wild type base A replaced by T"
FT   mutation        replace(478,G)
FT                   /*tag= l
FT                   /note= "Wild type base A replaced by G"
FT   mutation        replace(689,A)
FT                   /*tag= m
FT                   /note= "Wild type base G replaced by T"
FT   mutation        replace(690,G)
FT                   /*tag= n
FT                   /note= "Wild type base A replaced by G"
FT   mutation        replace(691,C)
FT                   /*tag= o
FT                   /note= "Wild type base G replaced by A"
FT   mutation        replace(728,C)
FT                   /*tag= p
FT                   /note= "Wild type base C replaced by G"
FT   mutation        replace(729,G)
FT                   /*tag= q
FT                   /note= "Wild type base C replaced by T"
FT   mutation        replace(730,G)
FT                   /*tag= r
FT                   /note= "Wild type base G replaced by A"
FT   mutation        replace(863,A)
FT                   /*tag= s
FT                   /note= "Wild type base G replaced by T"
FT   mutation        replace(864,A)
FT                   /*tag= t
FT                   /note= "Wild type base A replaced by G"
FT   mutation        replace(1016,T)
FT                   /*tag= u
FT                   /note= "Wild type base T replaced by G"
XX
PN   US20050281792-A1.
```

```
PD   22-DEC-2005.
XX
PF   01-SEP-2004; 2004US-00933115.
XX
PR   13-AUG-1997;   97US-00910798.
PR   01-MAR-1999;   99US-00259214.
PR   13-APR-1999;   99US-00291931.
PR   25-MAY-1999;   99US-00318528.
PR   25-MAY-2000;   2000US-00580515.
PR   24-MAY-2001;   2001US-00866379.
XX
PA   (SHOR/)  SHORT J M.
PA   (KRET/)  KRETZ K A.
PA   (GRAY/)  GRAY K A.
PA   (BART/)  BARTON N R.
PA   (GARR/)  GARRETT J B.
PA   (ODON/)  ODONOGHUE E.
PA   (BAUM/)  BAUM W.
PA   (ROBE/)  ROBERTSON D E.
PA   (ZORN/)  ZORNER P.
XX
PI   SHORT JM, Kretz KA, Gray KA, Barton NR, Garrett JB, Odonoghue E;
PI   Baum W, Robertson DE, Zorner P;
XX
PS   Claim 1; Page; 82pp; English.
XX
DR   WPI; 2006-0550022/06.
DR   P-PSDB; AEE75421.
XX
CC   The present sequence is that of a polynucleotide encoding a novel
CC   modified phytase derived from the Escherichia coli appA protein, with
CC   improved thermal tolerance and protease stability compared to the wild-
CC   type. The present invention relates to a novel formulation, comprising at
CC   least one polypeptide having phytase activity, where the polypeptide is a
CC   modified sequence (AEE75421) derived from the Escherichia coli K-12 appA
CC   phytase (AEE75419). The modification of the enzyme, comprising a series
CC   of mutations, improves the thermal tolerance and protease stability of
CC   the protein. The specification also claims a pharmaceutical composition;
CC   a kit; an immobilized phytase; a fertilizer or soil additive; a liquid
CC   supplement for preventing muscle cramps; a hydrating agent; a tissue
CC   culture or cell culture media; and a plant food additive; all comprising
CC   the novel phytase, and a method of reducing pollution and increasing
CC   nutrient availability in an environment or environmental sample by
CC   degrading environmental phytic acid. The novel enzyme is a dietary
CC   supplement useful for treating, preventing or reversing osteoporosis or
CC   bone loss, and preventing muscle cramps. The liquid supplement is useful
CC   for preventing muscle cramps. The method is useful for reducing pollution
CC   and increasing nutrient availability in an environment or environmental
CC   sample by degrading environmental phytic acid, where the environment or
CC   environmental sample comprises a soil or a body of water. The immobilized
CC   phytase is useful in foodstuffs for improving the feeding value of
CC   phytate rich ingredients. NOTE: This sequence is not shown in the
CC   specification but was created from the wild-type sequence (AEE75418) and
CC   the information given in claim 1.
XX
SQ   Sequence 1901 BP; 470 A; 497 C; 500 G; 433 T; 0 U; 1 Other;
```

```
Query Match             98.9%;  Score 1294.2;  DB 15;  Length 1901;
Best Local Similarity   99.7%;
Matches 1296;  Conservative      0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy           1 ATGAAAACGGATCTTAATCCCATTTTTATCCTTCTGATTCCGTTAACCCCGCAAATCTGCA 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         188 ATGAAAACGGATCTTAATCCCATTTTTATCCTTCTGATTCCGTTAACCCCGCAAATCTGCA 247

Qy          61 TTCGCTCAGAGTGAGCCCGGAGCTGGAAGTGTGGGTTGGGTTGAGTTGTCAGTCGTCATGGT 120
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         248 TTCGCTCAGAGTGAGCCCGGAGCTGGAAGTGTGGGTTGGGTTGAGTTGTCAGTCGTCATGGT 307
```

```
Qy    121 GTGGCGTGCTTCCAACCATCAAGAGGCCACGCAAACTGATTGCAGGATGTCACCCCAGACGCATGGCCA 180
          ||||| |||| |||||||          |||| ||| |||||| |||||||||||||||||||||||||
Db    308 GTGCCGTTGCTCCAACCAAGGCCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 367

Qy    181 ACCTGGCCGGTAAAAACTGGGTGAGCTGAACGCACCGCGGGGTGTGAGCTGAATCGCCTATCTC 240
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    368 ACCTGGCCGGTAAAAACTGGGTGAGCTGAACGCACCGCGGGGTGTGAGCTGAATCGCCTATCTC 427

Qy    241 GGACAATTACTGGCGTCAGCGTCTCGGTTAGCCGACGGATTGCTGCCTTAAAGTGGGCTGCCCG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428 GGACAATTACTGGCGGCCAGCGTCTTCGGTTAGCCGACGGATTGCTGCCTTAAAGTGGGCTGCCCG 487

Qy    301 CAGTCTGGTCAGGTGCGCAGATTATTGCTGTGAGTGCGACGGATTGCTGGTTAATCGGATCTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGGAGGCTACCCGTAAAACAGGCGAA 547

Qy    361 GCCTTCGCGCCCGCGGCGGGTGGGCCACCTGACTGTCAATAACCGTACATACCCAGGCACGATACG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548 GCCTTCGCGCCCGCGGCGGGTGGCCACCTGACTGTCAATAACCGTACATACCCAGGCACGATACG 607

Qy    421 TCCAGTCCCGGATCCGTTGTATTTAATCCTCTAAAAACTGGGCGTTTGCCAACTGGATAACGCG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608 TCCAGTCCCGGATCCGTTGTATTTAATCCTCTAAAAACTGGGCGTTTGCCAACTGGATAACGCG 667

Qy    481 AACGTGACTGACGCCGATCCCTCGAGGAGGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    668 AACGTGACTGACGCCGATCCCTCGAGGAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 727

Qy    541 TATCAAAACGGCGGTTTCGGGCGAACTGGAACCGGTTGCTTAAATTTTCCGCAATCAAAACTTGTGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    728 TATCAAAACGGCGGTTTCGGGCGAACTGGAACGGTTGCTTAATTTTCCGCAATCAAAACTTGTGC 787

Qy    601 CTTAAACGTGAGAAAACAGGAGCGAAAAGCTGTTCATTAACGAGCAGTCATTACCATCGGGAACTC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    788 CTTAAACGTGAGAAAACAGGAGCGAAAAGCTGTTCATTAACGAGCAGTCATTACCATCGGGAACTC 847

Qy    661 AAGGTGAGCGCCGCGACTGTCTCATTAACCTGGTGGGGTAAGCTTGCTCATCAAATGCTGACG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    848 AAGGTGAGCGCCGCGACTGTTGTCTCATTAACCTGGTAAGCCTCGCATCAAATGCTGACG 907

Qy    721 GAGATATTTCTCTGCAACAAGCACAGGGAATGCCGGAGCGGGGAAGGGAGGATCACC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908 GAGATATTTCTCTGCAACAAGCACAGGGAATGCCGGAGCGGGGAAGGGAGGATCACC 967

Qy    781 GATTCACACAGTGAGAACACACCTTGCTGCTTAGTTGTGTTTGCATAACGGCGGCGAAGAACTC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968 GATTCACACCAGTGGAACACCCTTGCTGCTTAGTTGCATAAAACGGGCAAATTGATTTGCTACAA 1027

Qy    841 CGCAGGCCCAGAGGTTGCCCCAGCGCCCAGCCCCCCGTTATTAGATTTGATCAAAGACAGCG 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028 CGCACGCCCAGAGGTTGCCCCGCAGCCCCGCTTATTTGAGTTGATCAAAGACAGCG 1087

Qy    901 TTGAGCGCCCCATCCACCGCAAAAACAGGGCTATGGCTGTGAGCTGACATTACCCACTTCAGTTGCTG 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 TTGAGCGCCCCATCCACCGCAAAAACAGGGCTATGGTGTGAGCTGACATTACCCACTTCAGTTGCTG 1147

Qy    961 TTTATCGCCGGACACGATACTAATCTGGGCAAAATCTCGGCGCGCACTGGAGGCTCAACTGG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148 TTTATCGCCGGACACGGATACTAATCTGGCAAAATCTCGGCGCGCACTGGAGCTCAACTGG 1207

Qy   1021 ACGCTTCCCCGGTCAGCCGGATAACACGCCCGGGATAACTGGTGTTTGAAGCCTGG 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208 ACGCTTCCCCGGTCAGCCGGATAACACGCCCGGGATAACTGGTGTTTGAACCCTGG 1267

Qy   1081 CGTGCGGCTAAGCGGATAACCAGTGGCGATTCAGGTTGGCTCTTCCAGACTTTACG 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268 CGTGCGGCTAAGCGGATAACAGCCAGTGGATCAGGTTTCGTCTGGCTCTTCAGACTTTACG 1327

Qy   1141 CAGATCGTGATAAAACGGCCCGCTGTCATTAAAATACCGCCGCTGTCATTAAAATACCGGACC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328 CAGATCGTGATAAAACGGCCCGCTGTCATTAAAATACCGCCGCTGTCATTAAAATACCGGACC 1387

Qy   1201 CTGGCAGGATGTGAAGAGCGGAAGGCGAAATGGGCAGGGGCCATGTTGGTTCGTTGGCCAGGGTTTTTAGGCAA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388 CTGGCAGGATGTGAAGAGCGGAAGGCGAAATGGGCAGGGGCCATGTTGGTTCGTTGGCCAGGGTTTTTAGGCAA 1447

Qy   1261 ATCGTGAATGAAGCACGCATACCGGGCGTGCAGTTTGAGAT 1300
          ||||||||||||||||||||||||||||||||||||| ||
Db   1448 ATCGTGAATGAAGCACGCATACCGGGCGTGCAGTTTGTAAT 1487
```

RESULT 6

Db ADO50296

ID   ADO50296 standard; DNA; 1901 BP.
XX
AC   ADO50296;
XX
DT   29-JUL-2004 (first entry)
XX
DE   Escherichia coli K12 appA phytase mutant DNA #2.
XX
KW   AppA phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW   phytate; animal feed; fish feed; dough; baking; gene; mutant; ds.
XX
OS   Escherichia coli; Kl2.
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   mutation   replace(437,C)
FT   /*tag= a
FT   mutation   replace(438,A)
FT   /*tag= b
FT   mutation   replace(439,A)
FT   /*tag= c
XX
PN   US20040091968-A1.
XX
PD   13-MAY-2004.
XX
PF   20-JUN-2003; 2003US-00601319.
XX
PR   13-AUG-1997; 97US-00910798.
PR   01-MAR-1999; 99US-00259214.
PR   13-APR-1999; 99US-00291931.
PR   25-MAY-1999; 99US-00318528.
PR   25-MAY-2000; 2000US-00580515.
PR   24-MAY-2001; 2001US-0086379.
XX
PA   (SHOR/) SHORT J M.
PA   (KRET/) KRETZ K.
PA   (GRAY/) GRAY K A.
PA   (BART/) BARTON N R.
PA   (GARR/) GARRETT J B.
PA   (ODON/) O'DONOGHUE E.
PA   (MATH/) MATHER E J.
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI   Mather EJ;
XX
DR   WPI; 2004-374952/35.
XX
PT   Producing phytase, involves providing nucleic acid derived from bacteria
PT   encoding polypeptide having phytase activity, and expressing nucleic acid
PT   in yeast.
XX
PS   Disclosure; SEQ ID NO 6; 74pp; English.
XX
CC   The invention relates to a method for producing phytase that involves
CC   providing a nucleic acid encoding phytase derived from a bacteria, and
CC   expressing the nucleic acid in a yeast under conditions that allow
CC   expression of the nucleic acid in the yeast. The invention also relates to
CC   modified phytase enzyme which has improved thermal tolerance and protease
CC   stability at low pH. The phytase enzyme can be used in foodstuffs to
CC   improve the feeding value of phytate rich ingredients, and in diet of
CC   numerous animals including mammals, fowls and fishes, commercially
CC   significant mammals such as pigs, goats, laboratory rodents, commercially

```
CC   significant avian species such as chicken, ducks, doves, parrot, etc.,
CC   commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC   etc., in dough making and baking, in dietary aids for animals. The method
CC   provides easy manufacture of the active ingredient loaded biocompatible
CC   composition, higher yields and loading efficiency. The phytase
CC   incorporated in to the dietary aid is safe for animals. The present
CC   sequence is Escherichia coli K12 appA phytase mutant DNA.
XX
SQ   Sequence 1901 BP; 472 A; 498 C; 501 G; 429 T; 0 U; 1 Other;

Query Match       97.0%;  Score 1268.6;  DB 12;  Length 1901;
Best Local Similarity   98.5%;  Pred. No. 0;  Mismatches 20;  DB 12;  Length 1901;
Matches 1280;  Conservative 0;  Mismatches 20;  Indels 0;  Gaps 0;

Qy      1 ATGAAAGGGATCTTAATCCCATTTTATCTCTTCTGATTCCGTTAACCCGCCAATCTGCA   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188 ATGAAAGGGATCTTAATCCCATTTTATCTCTTCTGATTCCGTTAACCCGCCAATCTGCA  247

Qy     61 TTCGCTCAGAGTGAGCCCGGAGCTGAAGCTGAAAAGTGTGGTGATCGTCGTTCATGGT  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248 TTCGCTCAGAGTGAGCCCGGAGCTGAAGCTGAAAAGTGTGGTGATCGTCGTTCATGGT  307

Qy    121 GTGGCGTGGTCCAACCAAGGCCACGCACGCAACTGATCAGCCAGGATGTCACCCCAGGCCA  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308 GTGGCGTGGTCCAACCAAGGCCACGCACGCAACTGATCAGCCAGGATGTCACCCCAGGCCA  367

Qy    181 ACCTGGCCGGGTAAAACTGGGGTGAGCTGATGCTAATCGCCTATCTC  240
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    368 ACCTGGCGGGGTAAAAACTGGGGTTGGCTGATGCTAATCGCCTATCTC  427

Qy    241 GGACATTACTGGCGTCAGCGTCTTCGGTTAGCCGACGGGATTGCTGCCTGCTGCTGCCCG  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428 GGACATTACTGGCGTCAGCGTCTTCGGTTAGCCGACGGGATTGCTGCCTGCTGCTGCCCG  487

Qy    301 CAGTCTCGGTCAGGTCGGCGGATTAATTGCTGAATGTCGACGAGCGGTACCCGTAAAACG  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488 CAGTCTCGGTCAGGTCGGCGGATTAATTGCTGAATGTCGACGAGCGGTACCCGTAAAACG  547

Qy    361 GCCTTCGCGCCCGGCCTGGCACCTGACTGTGCAATAACCGTACCATACCCCAGGCAGATACG  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548 GCCTTCGCGCCCGGCCTGGCACCTGACTGTGCAATAACCGTACCATACCCCAGGCAGATACG  607

Qy    421 TCCAGTCCCGACCCGTTATTAATCCCTCTAAAAAACTGGCGTTGCCAACTGGATAACGGCG  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608 TCCAGTCCCGATCCGTTATTAATCCCTCTAAAAAACTGGCGTTGCCAACTGGATAACGGCG  667

Qy    481 AACGTGACTGACGCGATCCTCGAGAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGGCAT  540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    668 AACGTGACTGACGCGATCCTCAGCGACGGGCAGGAGGGTCAAATTGCTGACTTTACCGGGGCAT  727

Qy    541 TATCAAAACGGCGCTTTCGCGACTGGAACGGGGTGCCTTAATTTTCCGCAATCAAACTTGTGC  600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    728 CGGCAAAACGGCGCTTTCGCGACTGGAACGGGGTGCCTTAATTTTCCGCAATCAAACTTGTGC  787

Qy    601 CTTAAACGTGAGAAACAGGACAGGAAGCTGTTCATTAAGCGCAGGCAGCTTAACCACCGGAACTC  660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    788 CTTAAACGTGAGAAACAGGACAGGAAGCTGTTCATTAAGCGCAGGCAGCTTAACCACCGGAACTC  847

Qy    661 AAGGTGAGCGCGGACACTGTGTCTCATTAACCGGTGCGGTAAGCCTGCGATCAATGCTGACG  720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    848 AAGGTGAGCGCGGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCGATCAATGCTGACG  907

Qy    721 GAGATATTTCTCTCGACACAGCACAGGGAATGCCGGAGCCCGGGTGGGGAGGGATCACC  780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    848 GAGATATTTCTCCTCGACACAGCACAGGGAATGCCGGAGCCCGGGTGGGGAGGGATCACC  967

Qy    781 GATTCACCACGTGGAACACTTGCTAAGTTTGCATAACGCGCCAAATTGATTTGCTACAA  840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968 GATTCACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAAATTTATTTGCTACAA  1027

Qy    841 CGCACGCCAGAGGTTGCCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG  900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028 CGCACGCCCAGAGGTTGCCCGCAGCCGCCACCCCGTTATTAGATTTGATCAAGAGACAGCG  1087

Qy    901 TTGACGCGCCCCATCCACCCGCAAAAACAGGGCGGTATGGTGGTGCGGGCGGTATACCCACTTCAGTGCTG  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 TTGACGCGCCCATCCACCCGCAAAAACAGGGCGGTATGGTGGTGCGGGCGGTATACCCACTTCAGTGCTG  1147

Qy    961 TTTATCGCCGGACACGATACTAATCTGGCGAAAAATCTCGGCGGCGGCGGCACTCGAGCTCAAACTGG  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148 TTTATCGCCGGACACGATACTAATCTGGCGAAAAATCTCGGCGGCGGCGGCACTCGAGCTCAAACTGG  1207

Qy   1021 ACGCGTTCCCGGTCAGCCCGGATAACAGCCGCGCCAGGTGGTGGAACTGGGTGTGTTGAACGCGCTGG  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208 ACGCGTTCCCGGTCAGCCCGGATAACAGCCGCGCCAGGTGGTGGAACTGGGTGTGTTGAACGCGCTGG  1267

Qy   1081 CGTCGGCTAAGCGGATAACAGCCAGTGGATTCAGTTTCGGCTGTCTTCCAGACTTTACAG  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268 CGTCGGCTAAGCGGATAACAGCCAGTGGATTCAGTTTCGGCTGTCTTCCAGACTTTACAG  1327

Qy   1141 CAGATGCCGTGATAAAACGCCGCTGTCTCATTAAAATACGCCGCCCGCCCGAGAGGGTGAAACTGACC  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328 CAGATGCCGTGATAAAACGCCGCTGTCTCATTAAAATACGCCGCCCGCCCGAGAGGGTGAAACTGACC  1387

Qy   1201 CTGGCAGATGTGAAAGAGCGGAAATGCGCAGGCCATGTGTTCGTTCGGTTGGGCAGGTTTTTACGCAA  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388 CTGGCAGATGTGAAAGAGCGGAAATGCGCAGGCCATGTGTTCGTTCGGTTGGGCAGGTTTTTACGCAA  1447

Qy   1261 ATCGTGAATGAAGCACGCATACCCGGCGTGCAGTTTGAGAT  1300
          |||||||||||||||||||||||||||||||||| ||||||
Db   1448 ATCGTGAATGAAGCACGCATACCCGGCGTGCAGTTTGTAAT  1487
```

RESULT 7
AD050295
ID   AD050295 standard; DNA; 1901 BP.
XX
AC   AD050295;
XX
DT   29-JUL-2004 (first entry)
XX
DE   Escherichia coli K12 appA phytase mutant DNA #1.
XX
KW   AppA phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW   phytate; animal feed; fish feed; dough; baking; gene; mutant; ds.
XX
OS   Escherichia coli, K12.
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   mutation   replace(389,T)
FT   /*tag= a
FT   mutation   replace(390,G)
FT   /*tag= b
XX
PN   US2004091968-A1.
XX
PD   13-MAY-2004.
XX
PF   20-JUN-2003; 2003US-00601319.
XX
PR   13-AUG-1997; 97US-00910798.
PR   01-MAR-1999; 99US-00259214.
PR   13-APR-1999; 99US-00291931.
PR   25-MAY-1999; 99US-00318528.
PR   25-MAY-2000; 2000US-00580515.
PR   24-MAY-2001; 2001US-00866379.
XX
PA   (SHOR/) SHORT J M.
PA   (KRET/) KRETZ K.
PA   (GRAY/) GRAY K A.
PA   (BART/) BARTON N R.
PA   (GARR/) GARRETT J B.
PA   (ODON/) O'DONOGHUE E.
PA   (MATH/) MATHER E J.

PI Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI Mather EJ;
XX
DR WPI; 2004-374952/35.
XX
PT Producing phytase, involves providing nucleic acid derived from bacteria
PT encoding polypeptide having phytase activity, and expressing nucleic acid
PT in yeast.
XX
PS Disclosure; SEQ ID NO 5; 74pp; English.
XX
CC The invention relates to a method for producing phytase that involves
CC providing a nucleic acid encoding phytase derived from a bacteria, and
CC expressing the nucleic acid in a yeast under conditions that allow
CC expression of the enzyme in the yeast. The invention also relates to
CC modified phytase enzyme which has improved thermal tolerance and protease
CC stability at low pH. The phytase enzyme can be used in foodstuffs to
CC improve the feeding value of phytate rich ingredients, and in diet of
CC numerous animals including mammals, fowls and fishes, commercially
CC significant avian species such as chicken, ducks, dove, parrot, etc.,
CC significant mammals such as pigs, goats, laboratory rodents, commercially
CC commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC etc. in dough making and baking, in dietary aids for animals. The method
CC provides easy manufacture of the active ingredient loaded biocompatible
CC composition, higher yields and loading efficiency. The phytase
CC incorporated in to the dietary aid is safe for animals. The present
CC sequence is Escherichia coli K12 appA phytase mutant DNA.
SQ Sequence 1901 BP; 475 A; 499 C; 499 G; 427 T; 0 U; 1 Other;

Query Match   96.9%;   Score 1267;   DB 12;   Length 1901;
Best Local Similarity   98.4%;   Pred. No. 0;   Mismatches 21;   Indels 0;   Gaps 0;
Matches 1279;   Conservative   0;   Mismatches   21;   Indels   0;   Gaps   0;

```
Qy      1  ATGAAAGCGGATCTTAATCCCATTTTTATTTTATCTCTGATTCCGTTAACCCCCCCAAATCTGCA   60
           ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    188  ATGAAAGCGGATCTTAATCCCATTTTTATTCTCTTCTGATTCCGTTAACCCCCCCAAATCTGCA  247

Qy     61  TTCGCTCGAGTGAGCCGAGGCTGAGCGTGGAAAGTGTGGTTCAGTTGTCGTCAGTGT        120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248  TTCGCTCAGAGTGAGCCGAGCTGAAGCTGAAAGTGTGGTTCAGTTGTCAGTGTCATGGT  307

Qy    121  GTGGCGTGCTCCAACAAAGGCCACGCAAACTGATGCAGGACATGTCACCCGAGGCCATGGCCA  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308  GTGGCGTGCTCCAACCAAGGCCACGCAAACTGATGCAGGACATGTCACCCCGAGACCATGGCCA  367

Qy    181  ACCTGCGCCGGTAAAAACTGGGTGAGCTGACACCGCGGGTGTGAGCTAATCGGCCTATCTC  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    368  ACCTGCGCCGGTAAAAACTGGGTGAGCTGACACCGCGGGTGTGAGCTAATCGGCCTATCTC  427

Qy    241  GGACATTACTGGCGTCAGCGTCTGGTGTTAGCCGACGGGATTGCTCCCTAAATGTNGGCTGCCG  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428  GGACATTACCAACGCCAGCGGTTCGGACTGCTGGGCGGAAAAGGGCTGCCG  487

Qy    301  CAGTCTGGTCAGGTCCCGATTATTGTCGATCTGACGCAGGCTACCCGGTAAAAACAGGCGAA  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488  CAGTCTGGTCAGGTCCCGATTATTGCTGATGTCGACGAGCGTACCCCGTAAAAACAGGCGAA  547

Qy    361  GCCTTCGGCCCGCGGCCGGCTGGCACCTGACTGTGCGAATAACCGTTACATACCAGGCAGATACG  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548  GCCTTCGGCCCGCGGCCGGCACCTGACTGTGCAATAACCGTTACATACCCAGGCAGATACG  607

Qy    421  TCCAGTCCCGATCCGGTTATTAATCCTCTAAAAACTGGCGTTGCCAACTGGATAACGCG  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608  TCCAGTCCCGATCCGGTTATTTAATCCTCTAAAAACTGGGCGTTGCCAACTGGATAACGCG  667

Qy    481  AACGTGACTGACGCGGATCGTCGAGAGGGCGCAGGAGGGTCAAATTGCTGACTTTACCGGCAT  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    668  AACGTGACTGACGCGGATCCTCAGCAGGGCAGGAGGGTCAAATGCTGACTTTACCGGCAT  727

Qy    541  TATCAAAACGGCGCGGTTTCGGGAACTGGAACGGGGTGCTTAATTTTTCGCAATCAAACTTGTGC  600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    728  CGGCAAACGGCGGTTTCGGGAACTGGAACGGGGTGCTTAATTTTTCGCAATCAAACTTGTGC  787

Qy    601  CTTAAACGTGAGAAAACAGGAGGAAAACAGTTCATTAACGCAGGCATTACCATCGGGAACTC  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    788  CTTAAACGTGAGAAAACAGGAGGAAAACAGTTCATTAACGCAGGCATTACCATCGGGAACTC  847

Qy    661  AAGGTGGAGCGGCCGGACTGTCTCATTAAACCGGTGCGGTCGTAAGCCTTCGGCATCGAATGCTGACG  720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    848  AAGGTGGAGCGGCGGACAAATGTCTCATTAAACCGGTGCGGTAAGCCTCCGGCATCGAATGCTGACG  907

Qy    721  GAGATATTTCTCCTGCAACAAGCACACAGGGAATGCCCGGAGCCGGGGGTTGGGGAGGGATACC  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908  GAGATATTTCTCCTGCAACAAGCACACAGGGAATGCCCGGAGCCGGGGGTTGGGGAGGGATACC  967

Qy    781  GATTCACACCAGTGGAACACCCTTGCTAAGTTTGCATAAACGGGCCAATTTGCTGATTTGCTCAA  840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968  GATTCACACCAGTGGAACACCCTTGCTAAGTTTGCATAAACGGGCCAATTTTATTGCTACAA  1027

Qy    841  CGCACGCGCCAGAGGTTGCCCGCGAGCCCCACCCCCGTTATTAGATTTGATTCAAGACAGCG  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028  CGCACGCGCCAGAGGTTGCCCGCGAGCCCCGGTTATTAGATTTGATTCAAGACAGCG  1087

Qy    901  TTGACCGCCCATCCACCGCAAAAACACAGGCGTATGGGGTGACATTACCCCACTTCAGTGCTG  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088  TTGACCGCCCATCCACCGCAAAAACAGGCGTATGGGGTGACATTACCCACTTCAGTGCTG  1147

Qy    961  TTTATCGCCCGGACACACGATACTAATCTCGGCGCGCCAAAATCTCGGCGGGCAAAATCTCACTGG  1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148  TTTATCGCCCGGACACGATACTAATCTCGGCGCGCCAAAATCTCGGCGGGCAAAATCTCACTGG  1207

Qy   1021  AGGCTTCCGGTCAGCGCCGGATAACACGCGCCCGCCCAGTGGGTGGTGGAACTGTGTTTGAACCGTG  1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208  AGGCTTCCGGTCAGCCGGATAACACGCGGGTGGTGGAACTGGTTGTTGAAACGGTGG  1267

Qy   1081  CGTCGGCTTAAGCGGATAACAGCCAGTGGATTCAGGTTTCGGCGTTCTTCCAGACTTTACAG  1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268  CGTCGGCTTAAGCGGATAACAGCCAGTGGATTCAGGTTTCGGCGTTCTTCCAGACTTTACAG  1327

Qy   1141  CAGATGGGTGATTAAAACGCCGCTGTCATTAAAATACGGGTGGATTCAGTTTCGCGAGAGGTGAAACTGACC  1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328  CAGATGGGTGATTAAAACGCCGCTGTCATTAAAATACGGGCCGCCCCGGGAGAGGTGAAACTGACC  1387

Qy   1201  CTGGCCAGGATGTGAAGAGCGGAAATGCCGAGGGCCAGTGGTGTTCGTTGGCAGGTTTACGGCAA  1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388  CTGGCAGGATGTGAAGAGCGGAAATGCCGAGGGCCAGTGGTGTTCGTTGGCAGGTTTACGGCAA  1447

Qy   1261  ATCGTGAATGAAGGCACACCCATACCGGGCTGCAGTTTGAGAT  1300
           ||||||||||||||||||||||||||||||||||||||||| ||
Db   1448  ATCGTGAATGAAGGACACCGCCATACCGGGCTGCAGTTTGTAAT  1487
```

```
PR    13-APR-1999;  99US-00291931.
PR    25-MAY-1999;  99US-00318528.
PR    25-MAY-2000;  2000US-00580515.
PR    24-MAY-2001;  2001US-00863379.
XX
PA    (SHOR/)  SHORT J M.
PA    (KRET/)  KRETZ K A.
PA    (GRAY/)  GRAY K A.
PA    (BART/)  BARTON N R.
PA    (GARR/)  GARRETT J B.
PA    (ODON/)  ODONOGHUE E.
PA    (BAUM/)  BAUM W.
PA    (ROBE/)  ROBERTSON D E.
PA    (ZORN/)  ZORNER P.
XX
PI    Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, Odonoghue E;
PI    Baum W, Robertson DE, Zorner P;
XX
DR    WPI; 2006-055022/06.
XX
PT    Formulation useful as dietary supplement for treating, preventing or
PT    reversing osteoporosis and bone loss, and preventing muscle cramps,
PT    comprises one or more polypeptides having phytase activity.
XX
PS    Disclosure; SEQ ID NO 5; 82pp; English.
XX
CC    The present sequence is a DNA sequence related to the production of the
CC    novel thermostable / protease resistant Escherichia coli phytase
CC    described in the specification. The present invention relates to a novel
CC    formulation, comprising at least one polypeptide having phytase activity,
CC    where the polypeptide is a modified sequence (AEE75421) derived from the
CC    Escherichia coli K-12 appA phytase (AEE75419). The modification of the
CC    enzyme, comprising a series of mutations, improves the thermal tolerance
CC    and protease stability of the protein. The specification also claims a
CC    pharmaceutical composition; a kit; an immobilized phytase; a fertilizer
CC    or soil additive; a liquid supplement for preventing muscle cramps; a
CC    hydrating agent; a tissue culture or cell culture media; and a plant food
CC    additive, all comprising the novel phytase, and a method of reducing
CC    pollution and increasing nutrient availability in an environment or
CC    environmental sample by degrading environmental phytic acid. The novel
CC    enzyme is a dietary supplement useful for treating, preventing or
CC    reversing osteoporosis or bone loss, and preventing muscle cramps. The
CC    liquid supplement is useful for preventing muscle cramps. The method is
CC    useful for reducing pollution and increasing nutrient availability in an
CC    environment or environmental sample by degrading environmental phytic
CC    acid, where the environment or environmental sample comprises a soil or a
CC    body of water. The immobilized phytase is useful in foodstuffs for
CC    improving the feeding value of phytate rich ingredients. NOTE: The
CC    present sequence is shown in the sequence listing but not discussed
CC    elsewhere in the specification.
XX
SQ    Sequence 1901 BP; 475 A; 499 C; 499 G; 427 T; 0 U; 1 Other;

Query Match             96.9%;  Score 1267;  DB 15;  Length 1901;
Best Local Similarity   98.4%;  Pred. No. 0;  Mismatches 21;  Indels 0;  Gaps 0;
Matches 1279;  Conservative   0;
```

```
Qy     1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCGCCAATCTGCA 60
           |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 ATGAAAGGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCGCCAATCTGCA 247

Qy    61 TTCGGTCAGAGTGAGCCGGAAGCTGAAAAGTGGTTGGTTGATTCAGTTCGTCATGGT 120
           |||| |||||||||||||||| |||||||||||| ||||| ||| ||||||||||||
Db   248 TTCGCTCAGAGTGAGCCGGAGCTGAAAAGTGTGGTTGATTCGGTTCAGTTCGTCATGGT 307

Qy   121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATAGTGCACCCCAGACGGCATGGCCA 180
           |||||||||||||||||||||||||||||||| |||||| |||| ||||||||||
Db   308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGGTGCACGCGGAAAGGCATGGCCA 367

Qy   181 ACCTCGGCCCGGTAAAACTGGGTGAGCTGACACCGGCCGGTAATGGCCTATCTC 240
           |||| ||||||||||||||||||| |||||||||||  ||  |||||||||||||
Db   368 ACCTGGCCCGGTAAAACTGGGTCAGCTGACACCGCGCGGNGGTGAGCTAATGGCCTATCTC 427

Qy   241 GGACATTACTGGGTCAGGTCTCAGGTAGCGGTAGCCGGACGGATTGCTGCCTAAAATGTGGCTGCCCG 300
           |||||||||  |   ||||  |  || ||||  || || |||||||||||||  ||  |  |||||||||
Db   428 GGACATTACCAACCGCAGGGTCTTGTAGCCGACGGATTGCTGGCGACGGATTGCCTAAAAAGGGCTGCCCG 487

Qy   301 CAGTCGGTCAGGTCGCGGATTATTGCTGATAGTGCGAACGAGCGGTACCGGTACCGGTAAAAAACAGGCGAA 360
Db   488 CAGTCGGTCAGGTCGCGGATTATTGCTGATAGTGCGAACGAGCGGTACCGGTAAAAAACAGGCGAA 547

Qy   361 GCCTTCGCGCCGCCGGCTGCCATTATTGCCGTGTCAATAACGTACTATCCCAGCGCAGATACG 420
Db   548 GCCTTCGCGCCGCCGGCTGGCAACTTGACGTGCAATAACCGTACATACCCAGCGCAGATACG 607

Qy   421 TCCAGTCCCGATCCGGTTATTTAATCCTCTAAAAACTGGGTTTTCCGCCAACTGGATAACGCG 480
Db   608 TCCAGTCCCGATCCGGTTATTTAATCCTCTAAAAACTGGGTTTTCCGCCAACTGGATAACGCG 667

Qy   481 AACGTGACTGACGCGGATCCTTCGAGAGGGGCGAGGAGGGTCAATTGCTGACTTTACCGGGCAT 540
Db   668 AACGTGACTGACGCGGATCCTTCAGCAGGGTCAATTGGCTGACTTTACCGGGCAT 727

Qy   541 TATCAAACGGGTTTTCGGAACTGGACGGGGTGCTTAAATTTCCGCAATCAAAACTTGTGC 600
Db   728 CGGCGAAACGGGGTTTTCGGAACTGGACGGGGTGCTTAAATTTCCGCAATCAAAACTTGTGC 787

Qy   601 CTTTAAACGTGAGAAAACAGGAGAAGCGTCTTCATTAACGCCAGGCCAATTACCGATCGGAACTC 660
Db   788 CTTTAAACGTGAGAAAACAGGAGAAGCGTCTTCATTAACGCCAGGCCAATTACCGATCGGAACTC 847

Qy   661 AAGGTGAGCGGCCGACTGTGTCTCATTAACCGGTGCGGTAAGCCTTCGCCATCAAATCGTGACG 720
Db   848 AAGGTGAGCGGCCGACTGTGTCTCATTAACCGGTGCGGTAAGCCTTCGCCATCAAATCGTGACG 907

Qy   721 GAGATATTTTCTCCTGCAACAAGCACACAGAGAAATGCCGGAGCCGGGGTGGGGAGGATCACC 780
Db   908 GAGATATTTTCTCCTGCAACAAGCACACAGAGAAATCCCGAGCCGGGGTGGGGAGGATCACC 967

Qy   781 GATTCACCAGTGGAACACCTTGCTAAGTTTGCATAAACGCGCAAATTGATTTGCTACAA 840
Db   968 GATTCACCACCAGTGGAACACCTTGCTAAGTTTGCATAAACGCGCAAATTTTATTTGCTACAA 1027

Qy   841 CGCACGCCAGAGGTTGCCCGCAGCCGGCCACCCGCCACCCCACCCGGTTATTAGATTTGATCAAGACAGCG 900
Db  1028 CGCACGCCAGAGGTTGCCCGCAGCCGGCCACCCGCCACCCCACCCGGTTATTAGATTTGATCAAGACAGCG 1087

Qy   901 TTGACGCCCATCCACCGCAAAAAACAGGCGTAATGGGTGACAATTACCCACTTCAGTGCTG 960
Db  1088 TTGACGCCCATCCACCGCAAAAAACAGGCGTAATGGGTGACAATTACCCACTTCAGTGCTG 1147

Qy   961 TTTATCGCGGACAGATACTAATCTGGCAAAATCTCGGCGCCCACTGGAGCTCAACTGG 1020
Db  1148 TTTATCGCGGACAGATACTAATCTGGCAAAATCTCGGCGCCCACTGGAGCTCAACTGG 1207

Qy  1021 ACGCTTCCCGGTCAGCCCGGATAAACACCCGCCAGGTGGTGAACTGGTTGAACGGTGG 1080
Db  1208 ACGCTTCCCGGTCAGCCCGGATAAACACCCGCCAGGTGGTGAACTGGTTGAACGGTGG 1267

Qy  1081 CGTCGGGCTAAGGGATAACAGCCAGTGGATTCAGGTTTCGCCTGGTCTTCCAGACTTTACAG 1140
Db  1268 CGTCGGGCTAAGGGGATAACAGCCAGTGGATTCAGGTTTCGCCTGGTCTTCCAGACTTTACAG 1327

Qy  1141 CAGATGCCTGATAAAAGCGCCCGCTGTCATTAAAATACGCGCCGGAGAGGTGAAACTGACC 1200
Db  1328 CAGATGCCTGATAAAAGCGCCCGCTGTCATTAAAATACGCGCCGGAGAGGTGAAACTGACC 1387

Qy  1201 CTGGCAGGATGTGAAGAGGCGAAATGCCCAGGCCATGTTGTTCGTTGGGCAGGGTTTACGGAA 1260
Db  1388 CTGGCAGGATGTGAAGAGGCGAAATGCCCAGGCCATGTTGTTCGTTGGGCAGGGTTTACGGCAA 1447

Qy  1261 ATCGGTGAATGAAAGAGCACCCATACCCGGCGGCTGCAGTTTGAGAT 1300
Db  1448 ATCGGTGAATGAAGAGCACCCATACCGGCGGCTGCAGTTTGTAAT 1487
```

```
RESULT 9
AAC88885
ID   AAC88885 standard; DNA; 1323 BP.
XX
AC   AAC88885;
XX
DT   07-MAR-2001 (first entry)
XX
DE   Escherichia coli B phytase enzyme nucleotide sequence.
XX
KW   Escherichia coli B; phytase enzyme; anabolic; phytate digestion;
KW   nutrition; ds.
XX
OS   Escherichia coli.
XX
PN   WO200071728-A1.
XX
PD   30-NOV-2000.
XX
PF   25-MAY-2000; 2000WO-US014846.
XX
PR   25-MAY-1999;  99US-00318528.
XX
PA   (DIVE-) DIVERSA CORP.
XX
PI   Short JM, Kretz KA;
XX
PT   Improving the nutritional value of phytate-containing foodstuffs, using
PT   phytase enzymes which catalyze the liberation of inorganic phosphate from
PT   the phytates.
XX
PS   Claim 2; Fig 1; 147pp; English.
XX
CC   The present sequence encodes a phytase enzyme from Escherichia coli B.
CC   The enzyme catalyses the liberation of inorganic phosphate from the
CC   phytate in phytate-containing foodstuffs and can thus be used to improve
CC   the nutritional value of phytate rich ingredients
XX
DR   WPI; 2001-112081/12.
DR   P-PSDB; AAB37892.
XX
SQ   Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;

Query Match             96.7%;  Score 1265;  DB 5;  Length 1323;
Best Local Similarity   98.0%;  Pred. No. 0;  Mismatches 26;
Matches 1280;  Conservative   0;  Mismatches 26;  Indels 0;  Gaps  0;

Qy          1 ATGAAAGCGATCTTAATCCTCTCATTTTTATCTCTTCGTTAACCCCCGCAATCTGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAAGCGATCTTAATCCATTTTTATCTCTTCGTTAACCCCCGCAATCTGCA 60

Qy         61 TTCGCTCAGAGTGACCCGGAGCTGAAAGTGTGTGATGATTGTCAGTCGTCATGGT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120

Qy        121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCGACGCCATGGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCGAGACGCCATGGCCA 180

Qy        181 ACCTGCCGCGGTAAAACTGGGTGAGCTGACACCCGCGGGGTTAATGCCTATTCTC 240
              |||||||||| ||||||||||| ||||||||||||||||||||||||||||||||||||
Db        181 ACCTGCCGCGGTAAAACTGGGTTGGCTGACACCCGCGNGGGTGGTTGGTTGGTTAATCGCCTATTCTC 240

Qy        241 GGACATTACTGGGCGTCAGCGTTCGGACGGATCGGGGTTGGCGTGCCCG 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGACAATTACCAAACGCCAGCGCGTCTCGGTTAGCCCGGACGGAAAAAACGGGCTGCCCCG 300

Qy        301 CAGTCTGGTCGGTCCGGATTATTGCTGATGAGTCGACGGAGCGCTACCGTGCCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGTCTGGTCAGGTGCGGGATTATTGCTGATGAGTCGACGGAGCGCTACCGTGCCGGAA 360

Qy        361 GCCTTCGCGCCGCGGCGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTTCGCGCCGCGGCGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420

Db        421 TCCAGTCCGGATCCGGTTATTTTAATCCTTCTAAAAACTGGGCGTTTGCCAACTGGATAACGCG 480

Qy        421 AACGTGACTGACGCGGATCCTCGAGAGGGGGCGCAGGAGGGTCAGTTGCTGGCGCAT 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACGTGACTGACGCGGATCCTCAGCAGGGGGCGCAGGAGGGTCAGTTGCTGGCGCAT 540

Qy        541 TATCAAACGCGGGTTTCGGCGAACTGGAACGGTGCTTAATTTTCCGCAATCAAAACTTGTGC 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGGCAAACGCGGGTTTCGGGCGAACTGGAACGGTGCTTAATTTTCCGCAATCAAAACTTGTGC 600

Qy        601 CTTAAACGTGAGAAACAGGACGAAAAGCTGTTCATTAACGCAGGCAGCATTACCATCGGAACTC 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTAAACGTGAGAAACAGGACGAAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC 660

Qy        661 AAGGTGAGGCGCCGACTGTGTCTCATTAACCGGTGCGGTGGCGTAAGCCTTGCCATCAAATGCTGACG 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGGTGAGGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTTCGCATCAAATGCTGACG 720

Qy        721 GAGATATTTCTCCTGCGAACAAGCACCAGCGGAATGCCGGACGCCGGGTTGGGGGGGGAAGGATCACC 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGATATTTCTCCTGCAACAAGCACAGCGGAATGCCGGACGCCGGGTTGGGGGGGAAGGATCACC 780

Qy        781 GATTCACACCAGTGGGGAACACCTTGCTAAGTTTGCATAAACGGGCAATTTGATTTGCTACAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAAACGGGCAATTTTATTTGCTACAA 840

Qy        841 CGCACGCCACCAGAGGTTGCCCCGGAGCGCCACCCCCGGTTATTGAGATCATCGGCAGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CGCACGCCAGAGGTTGCCCCGGAGCGCCACCCCCGGTTATTGATCATCGGCAGG 900

Qy        901 TTGACGGCCCCATCCACGGCAAAAAACAGGGCGTATAGTGTGTGACATTACCCACTTCAGTGCTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTGACGCCCCCATCCACGCAAAAAACAGGGCGTATGGTGTGTGACATTACCACTTCAGTACTG 960

Qy        961 TTTATCCGGACACGATACTAATCTGGGCAAAATCTCGGGCACGCTGGAGCTCAACTGG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTTATTGCCGGACACGGATACTAATCTGGGCAAAATCTCGGGCACGCTGGAGCTCAACTGG 1020

Qy       1021 ACGGCTTCCCGGGTTCAGCCGGGATAACACAGCCAGGTGGTGGGTGGGGTTGAACCGCTGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACGGCTTCCCGGGTTCAGCCGGGATAACACGCCCAGGGTGGTGAACACTGGGTGTTTGAAACGCTGG 1080

Qy       1081 CGTCCGGCTAGCGGATAACAGGCCAGTGGATTCAGGTTTCGCGCTGGCGCTGCTCCCGGCAATCTTTACAG 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CGTCCGGCTAGCGGATAACAGGCCAGTGGATTCAGGTTTCGCGCTGGCGCTGCTCCAGACTTTACAG 1140

Qy       1141 CAGATGCGGTGATAAAACGCCGCTGTCATTAAAAATACGGGCGCCCGGAGAGGTGAAACTGACC 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CAGATGCGGTGATAAAACGCCGCTGTCATTAAAAATACGGGCGCCCGGAGAGGTGAAACTGACC 1200

Qy       1201 CTGGCAGGATGTGAAAGAGCGAAAATGGGCAGGGGCATGTTCGTTGGCAGGGTTTTACGGCAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CTGGCAGGATGTGAAAGAGCGAAATGCCAGGGGCATGTTCGTTGGCAGGGTTTTACGGCAA 1260

Qy       1261 ATCGTGAAATGAAGCACGCCATACCGGGGGGTGCGCAGTTTGAGATCTCATC 1306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ATCGTGAAATGAAGCACGCCATACCGGGGGGTGCGCAGTTTGAGATCTCATC 1306
```

```
RESULT 10
AAD36473
ID   AAD36473 standard; DNA; 1323 BP.
XX
AC   AAD36473;
XX
DT   21-AUG-2002 (first entry)
XX
DE   Escherichia coli phytase DNA.
XX
```

```
KW   Dietary aid; biocompatible composition; therapeutic; digestive tract;
KW   foodstuff; digestion; phytase; enzyme; gene; ds.
XX
OS   Escherichia coli.
XX
FH   Key             Location/Qualifiers
FT
FT   CDS             1..1323
FT                   /*tag= a
FT                   /product= "E. coli phytase protein"
XX
PN   WO200189317-A2.
XX
PD   29-NOV-2001.
XX
PF   15-MAY-2001; 2001WO-US015764.
XX
PR   25-MAY-2000; 2000US-00580937.
XX
PA   (DIVE-) DIVERSA CORP.
XX
PI   Short JM,  Kretz KA,  O'donoghue B;
XX
DR   WPI; 2002-164149/21.
DR   P-PSDB; AAE22836.
XX
PT   New dietary aids comprising sustained release biocompatible compositions,
PT   comprise agent that assists in digestion, useful for delivering enzymes,
PT   therapeutics, medicine or agents to an organism.
XX
PS   Disclosure; Fig 1; 89pp; English.
XX
CC   The present invention relates to novel dietary aids comprising sustained
CC   release biocompatible composition which comprises an agent (enzymes such
CC   as phytase, amylase, esterase, protease) that assists in digestion. The
CC   biocompatible composition is effective upon oral consumption and release
CC   in the digestive tract of a subject. The dietary aids are useful for
CC   delivering enzymes, therapeutics, medicine and agents to an organism. The
CC   use of enzymes and other agents in digestive aids of livestock or
CC   domesticated animals not only improves the animal's health and life
CC   expectancy but also assists in increasing the health of livestock or in
CC   the production of foodstuffs from livestock. The present sequence is
CC   Escherichia coli phytase DNA
XX
SQ   Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;

     Query Match             96.7%;  Score 1265;  DB 6;  Length 1323;
     Best Local Similarity   98.0%;  Pred. No. 0;
     Matches 1280;  Conservative  0;  Mismatches  26;  Indels  0;  Gaps  0;

Qy          1 ATGAAAGCGATCTAAATCCCATTTTATCTCTTCTGATTCCGTTAACCCGCCAATCTGCA   60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAAGCGATCTAAATCCCATTTTATCTCTTCTGATTCCGTTAACCCGCCAATCTGCA   60

Qy         61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTCATTCGTCGTCATCATGGT  120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGCTGGAAAGTGTGGTGATTGTTCAGTCGTCATGGT  120

Qy        121 GTGCGTGCCTCCAACCAAGGCCACGCAACTGATGATGCAGGAGTGTCACCCGACGCATGGCCA  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGATGCAGGAGTGTCACCCAGAGGCATGGCCA  180

Qy        181 ACCTGGCCGGGTAAAACTGGGGTGAGCTGAACACCGACACCGCGCGGCGGTAATCGCCTATCTC  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCTGGCCGGGTAAAACTGGGGTGAGCTGAACACCGCCGCGCGGCGGTAATCGCCTATCTC  240

Qy        241 GGACATTACCAACGCAGCCGCTCGGCTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGACATTACCAACGCAGCCGCTCGGCTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG  300

Qy        301 CAGTCTCGGTCAGGTCCGGATTATTGCTGAGTCGGAGCGTACCCGTAAAAACAGGCGAA  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGTCTCGGTCAGGTCCGGATTATTGCTGAGTCGGACGAGCGTACCCGTAAAAACAGGCGAA  360

Qy        361 GCCTTCGCCGCCGCCGGCGTCGGCCACCTGTCGCAATAAACCGTACATACCCAGGCAGGATACG  420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTTCGCCGCCGCCGGCTGGCCACCTGTCGCAATAAACCGTACATACCCAGGCAGGATACG  420

Qy        421 TCCAGTCCCGGATCCGTTATTTAATCCTCTAAAAAACTGGGGTTTCCAACTGGATAAACGCG  480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCAGTCCCGGATCCGTTATTTAATCCTCTAAAAAACTGGGGCTTTCCAACTGGATAAACGTGC  480

Qy        481 AACGTGACTGACCGGGATCCTTCGAGAGGGCAGGAGGAGGTCAATTTGCTTACCGGGGCAT  540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACGTGACTGACCGGGATCCTTCGAGAGGGGCAGGAGGAGGTCAATTTGCTTACCGGGCAT  540

Qy        541 TATCAAAACGGCGGTTTCGGGAAACTGGAACGGGGTGCTTGCTTAATTTCCGCAATCAACTTGTGC  600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGGCAAACGGCGGTTTCGGCGGAACTGGAACGGGGTGCTTAATTTCCGCAATCAAACTTGTGC  600

Qy        601 CTTAAAACGTGAGAAAACAGGAGACGAAAAGCTTGTTCATTACGCCAGGGCATTACCATCGGAACTC  660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTAAAACGTGAGAAAACAGGAGACGAAAAGCTTGTTCATTACGCCAGGGCATTACCATCGGAACTC  660

Qy        661 AAGGTGAGCGCCCGACTGTGTCTCATTAACCGGTGCGGGTAAGCCTTCGCATCAACAATGCTGACG  720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGGTGAGCGCCCGACAATGTCTCATTAACCGGTGCGGGTAAGCCTTCGCATCAAACAATGCTGACG  720

Qy        721 GAGATATTTCTCCTTGCAACAAGCACAGGGAATGCCGGAGCCGGGGGTGGGGAAGGGATCACC  780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGATATTTCTCCTGCAACAAGCACAGGGAAATGCCGGAGCCGGGGGTGGGGAAGGGATCACC  780

Qy        781 GATTCACACCAGTGGAACACCTTGCTCAAGTTTGCATAACGGCGCAATTTTATTTGGTCTACAA  840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAATTCACACCAGTCGGAACACCTTGCTCAAGTTTGCATAACGGCGCAATTTTATTTGGTCTACAA  840

Qy        841 CGCACGGCCAGAGCGGTTGCCCGGCACCCCGCCACCCCCGTTATTGAGATTTGATCAAGACAGCG  900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CGCACGGCCAGAGCGGTTGCCCGGCACCCCGCCACCCCCGTTATTGAGATTGATCATGGCACAGCG  900

Qy        901 TTGACGCCCCATCCACCGCAAAACAGGGCGTATGGGTTGACATTACCCAACTTCAGTCGCTG  960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGGTTGACATTACCCACTTCAGTACTG  960

Qy        961 TTTATTGCCCGGACACGATACTAATCCTGGCAAAATCTCGGCGGCCGCGGCCGACTGGAGCTCAACTGG  1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTTATTGCCGGACACGATACTAATCCTGGCAAAATCTCGGCGGCCGCGGCCGACTGGAGCTCAACTGG  1020

Qy       1021 ACGCTTCCCGGTCAGCCGGATAACACCGCGCGCCAGGGTGGTGAACTGGGTGTTGAACGCTGG  1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACCGTTCCCGGTCAGCCGGATAACACCGCGCCGCCAGGGTGGTGAACTGGGTGTTGAACGCTGG  1080

Qy       1081 CGTCGGCGTAAGCCGGATAACACCAGCGCCAGTGGAATTCAGGTTTCGCGTCTTCCAGACTTTACAG  1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CGTCGGCGTAAGCCGGATAACAGCCAGTGGAATTCAGGTTTCGCGTCTTCCAGACTTTACAG  1140

Qy       1141 CAGATGGGTGATAAAAACGCGCTGCTGTCATTAAATACGCGCCCGGCGAGAGGTGAAACTGACC  1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CAGATGGGTGATAAAAACGCGCTGCTGTCATTAAATACGCGCCCGGCGAGAGGTGAAACTGACC  1200

Qy       1201 CTGGCAGGATGTGAAAGAGGCGGAAATGCCGGAGCCGATGTGTTCGGTTCGTTCGCAGGTTTTACGGCCA  1260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CTGGCAGGATGTGAAAGAGGCGGAAATGCCGGAGCCGATGTGTTCGGTTCGTTCGCAGGTTTTACGGCCAA  1260

Qy       1261 ATCGTGAATGAAGCACGGCATACCGGCGGTGCAGTTTGAGATCTCATC  1306
              |||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ATCGTGAATGAAGCACGGCATACCGGCGGTGCAGTTTGAGATCTCATC  1306
```

RESULT 11
AAD25460
ID   AAD25460 standard; DNA; 1323 BP.
XX
AC   AAD25460;
XX
DT   26-MAR-2002 (first entry)
XX

```
DE   Escherichia coli B phytase DNA.
XX
KW   Bacterial phytase; K12 appA phytase; protease stability; anabolic;
KW   gastrointestinal; nutritional value; feed treatment process; therapy;
KW   thermal tolerance; growth performance; alcoholic drink; biobleaching;
KW   non-alcoholic drink; biobleaching; B phytase; ds.
XX
OS   Escherichia coli.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..1323
FT                  /*tag= a
FT                  /product= "E. coli B phytase protein"
XX
PN   WO20019333-A2.
XX
PD   29-NOV-2001.
XX
PF   24-MAY-2001; 2001WO-US017118.
PF   25-MAY-2000; 2000US-00580515.
XX
PR   25-MAY-2000; 2000US-00580515.
XX
PA   (DIVE-) DIVERSA CORP.
XX
PI   Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, O'donoghue E;
XX
DR   WPI; 2002-083108/11.
DR   P-PSDB; AAE15806.
XX
PT   New bacterial phytase for e.g. improving the nutritional value of phytate
PT   -containing foodstuffs and subsequently improving the growth performance
PT   of an organism that consumes it, or in treating animal digestive systems.
XX
PS   Claim 2; Fig 1; 170pp; English.
XX
CC   The patent discloses recombinant bacterial phytase from Escherichia coli
CC   K12 appA phytase. The enzyme has phytase activity and improved thermal
CC   tolerance when compared with wild-type phytase. It has improved protease
CC   stability at low pH. The recombinant phytase is useful for improving the
CC   nutritional value of phytate-containing foodstuffs and subsequently
CC   improving the growth performance of an organism that consumes it, in
CC   treating animal digestive systems, in feed treatment processes and for in
CC   vitro purposes related to research, discovery and development. They are
CC   also used for generating recombinant digestive system life forms, for
CC   producing or manufacturing alcoholic and non-alcoholic drinks based on
CC   the use of moulds, grains and/or plants, in biopulping and bio- bleaching
CC   where a reduction in the use of environmentally harmful chemicals that
CC   are traditionally used in the pulp and paper industry is desired and in
CC   the reduction or possible elimination of the need for mineral
CC   supplements, enzymes or therapeutic drugs for animals present in the feed
CC   thus increasing the amount calories and nutrients present in the daily feed.
CC   The present sequence is a DNA encoding E. coli B phytase protein
XX
SQ   Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;

Query Match       96.7%;   Score 1265;   DB 6;   Length 1323;
Best Local Similarity  98.0%;   Pred. No. 0;
Matches  1280;   Conservative  0;   Mismatches  26;   Indels  0;   Gaps  0;

Qy     1  ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCCTTCGATTCCGTTAACCCCGCAATCTGCA   60
          ||||||||||||||||||||||||||||  | |||||||||||||||||||||||||||||||
Db     1  ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCCTTCGATTCCGTTAACCCCGCAATCTGCA   60

Qy    61  TTCGCTCAGAGTGAGCCGGAGCTGAAGCTCGAAAAGTGTGGTGGATGTTCTCAGTCGTCATGGT  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGTGGTGATGTTCTCAGTCGTCATGGT  120

Qy   121  GTGCGTGCTTCCAACCAAGGCCACGACGCAACTGATGCAGGAGTCACCCGAGACCCATGGCCA  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGAAATGGGCGAAAATGGGCGCCA  180

Qy   181  ACCTGGCCGGGTTAAAACTGGGGTCAGCTAACACCGCGCCGGTGGTTGAGCTTAATGCCTATCTC  240
          ||||||||||||||||||||||||||||||||||||||  | ||||||||||||||||||||||
Db   181  ACCTGGCCGGGTTAAAACTGGGGTCAGCTAACACCGCGCGNGGTGGTTGAGCTTAATGCCTATCTC  240

Qy   241  GGACATTACTGGCGTCAGGTCTGGGTAGCCGACGGATTGCTCCTTAAATGTGGCTGCCG  300
          ||||||||||  | |||||||||||||||||||||||| ||||  | |||||||||||
Db   241  GGACATTACCAACGCCAGCGGTCGGATAGCCGACGGATTGCTGGGAAAAAGGGCTGCCG  300

Qy   301  CAGTCTGGTCAGGTCCGCATTATTTGCTGATGATGTCGAGACGGCGTGACCCGTAAAACAGGCGAA  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  CAGTCTGGTCAGGTCCGCATTATTTGCTGATGTCGAGACGGCGTTACCCGTAAAACAGGCGAA  360

Qy   361  GCCTTCGCGCCGCCGGGGGCTGGGCCACCTGACCTTGCTGTGCCAATAACCGTACATACCGAGATACG  420
          ||||||||||||  | |||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  GCCTTCGCGCCGCGGGGGCTGGGCCACCTGACCTTGCTGTGCCAATAACCGTACATACCGGATAACG  420

Qy   421  TCCAGTCCCGGATCCGTTATTTAATCCTTCTAAAAACTGGCGTTGCCAACTGGATAACGCG  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  TCCAGTCCCGGATCCGTTATTTAATCCTTCTAAAAACTGGCGTTGCCAACTGGATAACGCAT  480

Qy   481  AACGTGACTGACGCGGATCCTCAGCAGGGCAGGAGGGTCAATTTGCTGACTTTACCGGGCAT  540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  AACGTGACTGACGCGGATCCTCAGCAGGGCAGGAGGGTCAATTTGCTGACTTTACCGGGCAT  540

Qy   541  TATCAAACGGCGTTTCGCGAACTGGAACGGGGTGCTTAAATTTCCGCAAATCAAACTTGTGC  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  CGGCAAACGGCGTTTCGCGAACTGGAACGGGGTGCTTAAATTTCCGCAATCAAAACTTGTGC  600

Qy   601  CTTAAACGTGAGAAACAGGACGAAAGTTGTTCATTAACGCAGCAGGCATTACCATCGGAACTC  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  CTTAAACGTGAGAAACAGGACGAAAGTTGTTCATTAACGCAGCAGGCATTACCATCGGAACTC  660

Qy   661  AAGGTGAGCGCCGACAAATGTCTCATTAACCCGTGCGGTAAGCCTCGCATCAATGCTGACG  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  AAGGTGAGCGCCGACAAATGTCTCATTAACCCGTGCGGTAAGCCTCGCATCAATGCTGACG  720

Qy   721  GAGATATTTCTCCTGCCAACAAGCCACAGGGAATGCCCGAGCCCGGGGTGGGGAAGGATCACC  780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  GAGATATTTCTCCTGCCAACAAGCACAGGGAAATGCCCGAGCCCGGGGTGGGGAAGGATCACC  780

Qy   781  GATTCACCAGCTGGAACACCTTGCTAAGTTTGCATAAACGGCGCCAATTGATTGCTACAA  840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  GATTCACCAGCTGGAACACCTTGCTAAGTTTGCATAAACGGCCAATTTATTGGCTACAA  840

Qy   841  CGCACCGCCAGAGGTTGCCCGGCAGCCGGGCCACCCCGGTTATTGGATTGATCACGGCAGCG  900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  CGCACCGCCGAGGTTGCCCGGCAGCCGGGCCACCCCGGTTATTGGATTGATCAGGCAGCG  900

Qy   901  TTGACGCCCATCCACGGCAAAAAACAGGCGTATGGTGACATTACCACTCAGTCCTG  960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  TTGACGCCCATCCACGGCAAAAACAGGCGTATGGTGACATTACCACTCAGTACTG  960

Qy   961  TTTATCGCGGACAACGATACTAATCTGGCAAAATCTCGGCGACTGGAGCTCAACTGG  1020
          ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  TTTATTGCGGACAACGATACTAATCTGGCAAATCTCGGCGACTGGAGCTCAACTGG  1020

Qy  1021  ACGGCTTCCCGGTCAGCGGGATAACAGCCCGCGCCAGGGTGGTGAACTGGGTGTTTGAAACGGTGG  1080
          ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  ACGGCTTCCGGTCAGCCGGATAACAGCCCGCGCCAGGGTGGTGAACTGGGTGTTTTGAAACGGTGG  1080

Qy  1081  CGTTCGGCTTAAGCGGATAACAGCCAGTGGAATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  CGTTCGGCTTAAGCGGATAACAGCCAGTGGAATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG  1140

Qy  1141  CAGATGCGTGATAAAAGCGGCTGTCATTAAAATACGGCGCCCGGAGGGTGAAAACTGACC  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  CAGATGCGTGATAAAAGCGGCTGTCATTAAAATACGGCGCCCGGAGAGGGTGAAAACTGACC  1200

Qy  1201  CTGGCAGGATGTGAAGAGGCGAAATGCGGCAGGTGAAAGGCCATGTTTGAGATCTCATC  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  CTGGCAGGATGTGAAGAGGCGAAATGCCCGCAGGTGAAAGGCCATGTTTGAGATCTCATC  1260

Qy  1261  ACCTGGCCGGGTTAAAACTGGGGTCAGCTAACACCGCGCCGGTGCAGTTTGAGATCTCATC  1306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  ATCGTGAATGAAGCACGCCATACCGGCGTGCCATGCCGGGGTGCAGTTTGAGATCTCATC  1306
```

RESULT 12

ID   ADA19445 standard; DNA; 1323 BP.
XX
AC   ADA19445;
XX
DT   20-NOV-2003 (first entry)
XX
DE   E. coli B DNA encoding phytase.
XX
KW   Phytase; ds; gene; phytate; animal feed; inorganic phosphate;
KW   digestion enhancement; transgenic; thermal tolerance; protease stability.
XX
OS   Escherichia coli B.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1323
FT   /tag= a
FT   /product= "Phytase"
XX
PN   US20021367541-A1.
XX
PD   26-SEP-2002.
XX
PF   24-MAY-2001; 200IUS-0086379.
XX
PR   13-AUG-1997; 97US-00910798.
PR   01-MAR-1999; 99US-00255914.
PR   13-APR-1999; 99US-00291931.
PR   25-MAY-2000; 99US-00318528.
PR   25-MAY-2000; 2000US-00580515.
XX
PA   (SHOR/) SHORT J M.
PA   (KRET/) KRETZ K A.
PA   (GRAY/) GRAY K A.
PA   (BART/) BARTON N R.
PA   (GARR/) GARRETT J B.
PA   (DONO/) O' DONOGHUE E.
PA   (MATH/) MATHUR E J.
XX
PI   Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB;
PI   O' Donoghue E, Mathur EJ;
XX
DR   WPI; 2003-040002/03.
DR   P-PSDB; ADA19446.
XX
PS   Claim 2; Fig 1; 62pp; English.
XX

CC   The invention relates to an isolated Escherichia coli polynucleotide
CC   encoding a phytase enzyme appearing as ADA19450 and having amino acids
CC   modification from W68S, Q84W, A95P, K97C, S168E, R180Y, N26C or Y277D.
CC   Also included the E. coli appA gene ADA19449 (or an oligonucleotide
CC   derived from it) or its mutant sequence ADA19452, expression vectors,
CC   host cells, a method of improving nutritional value of a phytate-
CC   containing foodstuff by contacting the phytate-containing foodstuff with
CC   a pure phytase enzyme mutated phytase (where the phytase enzyme catalyzes
CC   the liberation of inorganic phosphate from the phytate in the phytate-
CC   containing foodstuff), a method to produce an animal feed containing a
CC   microbial phytase (comprising culturing the plant cell, plant part or
CC   plant under conditions where the nucleotide sequence is expressed and
CC   converting the plant cells, plants or plant into a composition for animal
CC   feed), a feed composition for animals (comprising the plant seeds, plant
CC   cells, plant parts or plants in admixture with a phytate-containing
CC   foodstuff), a method to treat a human or an animal able to benefit from
CC   digestive enhancement by the activity of an exogenous phytase enzymes
CC   comprising administering to the human or animal the plant seed, plant

CC   cells, plant parts or plants of a transgenic plant which is modified to
CC   contain an expression system which expresses a nucleotide sequence
CC   encoding a phytase enzyme, a transgenic non-human organism whose genome
CC   comprising a heterologous nucleic acid sequence encoding a polypeptide
CC   having phytase activity. The phytase enzyme is useful for improving the
CC   nutritional value of phytate-containing foodstuff, in the production of
CC   animal feed and for enhancing digestion in humans and animals. The
CC   invented method improves thermal tolerance and protease stability. It
CC   also improves the feeding value of phytate rich ingredients. The present
CC   sequence encodes E. coli B wild-type phytase.
XX
SQ   Sequence 1323 BP; 323 A; 353 C; 289 T; 0 U; 1 Other;

Query Match   96.7%; Score 1265; DB 8; Length 1323;
Best Local Similarity   98.0%; Pred. No. 0;
Matches 1280; Conservative   0; Mismatches   26; Indels   0; Gaps   0;

```
Qy          1 ATGAAAGCGATCTTAATCCCATTTTATCTTCTTCTGATTCCGGTTAACCCCGCAATCTGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAAGCGATCTTAATCCCATTTTATCTTCTTCTGATTCCGGTTAACCCCGCAATCTGCA 60

Qy         61 TTCGCCTCAGAGTGAGGCCCGGAGCTGGAAAGTGTGGGTGATTGTCAGTCGTCGTCATGGT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGCCTCAGAGTGAGGCCCGGAGCTGGAAAGTGTGGGTGATTGTCAGTCGTCGTCATGGT 120

Qy        121 GTGGGTGCTCCAACCAAGGCCAGCCAACTGATGCGAGGAGTGTCACCCCAGACGCATGGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTGGGTGCTCCAACCAAGGCCAGCCAACTGATGCGAGGAGTGTCACCCCAGACGCATGGCCA 180

Qy        181 ACCTGGGCCGGTAAAACTGGGTGAGCTGACCACCCGCGCGTTGGGTTGAGCTAATCGCTATCTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCTGGGCCGGTAAAACTGGGTTGGCTGACCACCCGCGGNGGGTGGAGCTAATCGCTATCTC 240

Qy        241 GGACATTACTGGGGCTGCACGTGACCGGTAGCCGACGGATTGCCTGCTTAAAGTGGGTGCCCG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGACATTACCAACGGCCAGCGTCTGGTAGCCGACGGATTGCTGGCGGAAAAAAGGGGTGCCCG 300

Qy        301 CAGTCTGGTCAGGTCGCGGATTATTGCTGCATGTGACGAGCGGTACCGGTAAAAACAGGGCGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGTCTGGTCAGGTCGCGGATTAATTGCTGCATGNTGACGAGCGGTACCGGTAAAAACAGGGCGAA 360

Qy        361 GCCTTCGCCGCCGCCGGGCTTGGCACCTGACTGTGCAATAACCGGTACACTACCCAGGCGAGGATACG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTTCGCCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGGTACACTACCCAGGCGAGGATACG 420

Qy        421 TCCAGTCCCGATCCGTTATTTAATTCCTCTAAAAACTGGGCGTTGCCAACTGGATAACGCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCAGTCCCGATCCGTTATTTAATTCCTCTAAAAACTGGGCGTTTGCCAACTGGATAACGCGG 480

Qy        481 AACGTGACTGACCGGATCCTCAGCAGGGCCAGAGGGGTCAATTGCTGACTTTACCGGGGCAT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACGTGACTGACCGGATCCTCAGCAGGGCCAGAGGGGTCAATTGCTGACTTTACCGGGGCAT 540

Qy        541 TATCAAACGGGCGTTTCCCGGAACGGGGTGCTTAAATTTTCCGCCAATCAAACTTGTGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGGCCAAACGGGCGTTTCCCGGAACGGGGTGCTTAAATTTTCCGGCAATCAAACTTGTGC 600

Qy        601 CTTAAACGTGAGAAAACAGGACGGAAAGCTGTTCATTAACGGCAGGCAATTACCATCGGAACTC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTAAACGTGAGAAAACAGGACGGAAAGCTGTTCATTAACGGCAGGCATCAAATGCTGACTC 660

Qy        661 AAGGTGAGCGCCGACACTGTGTCTCATTAACGGTGCGCGTAAGCCTGACG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGGTGAGCGCCGACAAGTCTCATTAACCGGTGCGCGTAAGCCTGACG 720

Qy        721 GAGATATTTCTCCTGCAACAGCACAGGGGAATGCCGGAGCCCGGCGTGGGGGGAGGATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGATATTTCTCCTGCAACAGCACAGGGGAATGCCGGAGCGTCGGGGGAGGGATCACC 780

Qy        781 GATTCACCAGGAACACCTTGCTAAGTTTGCATAACGCGCAAATTTTATTGCTACAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAAATTTTATTGCTACAA 840
```

```
Qy    841  CGCACGCCAGAGGTTGCCCCCAGCCGCCGCCACCCCGGTTATTATTAGATTTGATCAAGACAGCG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  CGCACGCCAGAGGTTGCCCCCAGCCGCCGCCACCCCGGTTATTATTAGATTTGATCAAGACAGCG  900

Qy    901  TTGACGCCCCATCCACCGCAAAAACAGGCCCGCGCCGCCACCCCGTTATTGGATTTGATCATGGCAGCG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  TTGACGCCCCATCCACCGCAAAAACAGGCCCGCGCCGCCACCCCGTTATTGGATTTGATCATGGCAGCG  960

Qy    961  TTTATGCCGGGACACGATACTACTTCGGCAAATCTTCGGGCGGCGCCACTGGGAGCTCAACTGG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  TTTATTGCCGGACACGATACTACTTCGGCAAATCTTCGGGCGGCGCCACTGGGAGCTCAACTGG  1020

Qy   1021  ACGCTTCCCGGTCAGCCGGATAACACGCCCCGGATAACACGCCCCAGTGTGTTGTTTGAACGGCTGG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  ACGCTTCCCGGTCAGCCGGATAACACGCCCCGGATAACACGCCCCAGTGTGTTGTTTGAACGGCTGG  1080

Qy   1081  CGTCCGGCTAAGCGGATAACACGCCAGTGGGATTCAGGTTCGCTGGGTCTTCCAGACTTTACAG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  CGTCCGGCTAAGCGGATAACACGCCAGTGGGATTCAGGTTCGCTGGGTCTTCCAGACTTTACAG  1140

Qy   1141  CAGATGCGTGATAAAACGCCCGCCTGTCATTAAATACGCCCGCGCCGGAGAGGTGAAACTGACC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  CAGATGCGTGATAAAACGCCCGCCTGTCATTAAATACGCCCGCGCCGGAGAGGTGAAACTGACC  1200

Qy   1201  CTGGCAGGATGTGAGAGAGCGAAAATCGCGCAGGGCATGTGTTCGTTGGCAGGTTTTTACGCAA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  CTGGCAGGATGTGAGAGAGCGAAAATGGCACAGGGCATGTGTTCGTTGGCAGGTTTTTACGCAA  1260

Qy   1261  ATCGGTGAATGAAGCACGCATACCGGCGGGTGCAGTTTGAGATCTCATC
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   1261  ATCGTGAATGAAGCACGCATACCGGCGGGTGCAGTTTGAGATCTCATC  1306
```

```
RESULT 13
AD050291

ID   AD050291   standard; DNA; 1323 BP.
XX
AC   AD050291;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Escherichia coli B phytase DNA.
XX
KW   Phytase; bacterial; thermal tolerance; protease stability; foodstuff;
KW   phytate; animal feed; fish feed; dough; baking; gene; ds.
XX
OS   Escherichia coli; B.
XX
FH   Key            Location/Qualifiers
FH
FT   CDS            1..1323
FT                  /*tag= a
FT                  /product= "Phytase enzyme"
FT   unsure         214..216
FT                  /*tag= b
FT                  /note= "Encodes Arg"
XX
PN   US200491968-A1.
XX
PD   13-MAY-2004.
XX
PF   20-JUN-2003; 2003US-00601319.
XX
PR   13-AUG-1997;  97US-00910798.
PR   01-MAR-1999;  99US-0029214.
PR   13-APR-1999;  99US-00291931.
PR   25-MAY-1999;  99US-00318528.
PR   25-MAY-2000;  2000US-00500515.
PR   24-MAY-2001;  2001US-0086379.
XX
PA   (SHOR/) SHORT J M.
PA   (KRET/) KRETZ K.
PA   (GRAY/) GRAY K A.
PA   (BART/) BARTON N R.
```

```
PA   (GARR/) GARRETT J B.
PA   (ODON/) O'DONOGHUE E.
PA   (MATH/) MATHER E J.
XX
PI   Short JM,   Kretz K,   Gray KA,   Barton NR,   Garrett JB,   O'donoghue E;
PI   Mather EJ;
XX
DR   WPI; 2004-374952/35.
DR   P-PSDB; AD050292.
XX
PS   Claim 3; SEQ ID NO 1; 74pp; English.
XX
CC   The invention relates to a method for producing phytase that involves
CC   providing a nucleic acid encoding phytase derived from a bacteria, and
CC   expressing the nucleic acid in a yeast under conditions that allow
CC   expression of the nucleic acid in the yeast. The invention also relates to
CC   modified phytase enzyme which has improved thermal tolerance and protease
CC   stability at low pH. The phytase enzyme can be used in foodstuffs and in
CC   improve the feeding value of phytate rich ingredients, and in diet of
CC   numerous animals including mammals, fowls and fishes, commercially
CC   significant mammals such as pigs, goats, laboratory rodents, commercially
CC   significant avian species such as chicken, ducks, doves, parrot, etc.,
CC   commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC   etc., in dough making and baking, in dietary aids for animals. The method
CC   provides easy manufacture of the active ingredient loaded biocompatible
CC   composition, higher yields and loading efficiency. The phytase
CC   incorporated in to the dietary aid is safe for animals. The present
CC   sequence is Escherichia coli B phytase DNA.
XX
SQ   Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;
```

```
Query Match   96.7%;  Score 1265;  DB 12;  Length 1323;
Best Local Similarity   98.0%;  Pred. No. 0;
Matches 1280;  Conservative    0;  Mismatches  26;  Indels   0;  Gaps   0;

Qy     1  ATGAAAGCGATCTTAATCCCACATTTTTATCTCTTCTGATTCGTTAACCCCCGCAAATCTGCA
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  ATGAAAGCGATCTTAATCCCACATTTTTATCTCTTCTGATTCGTTAACCCCCGCAAATCTGCA   60

Qy    61  TTCGCTCGAGAGTGGAGCCGGAGCTGAAGCTGAAAAGTGTGGTGGAATTGTCAGTCGTCATGGT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  TTCGCTCGAGAGTGGAGCCGGAGCTGAAGCTGAAAAGTGTGGTGGAATTGTCAGTCGTCATGGT  120

Qy   121  GTGCGTGCCTCCAACCAAGGCCACGCAACTGATTGCAGGATGTCACCCCGAGACGACCATGGCCA
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GTGCGTGCCTCCAACCAAGGCCACGCAACTGATTGCAGGATGTCACCCCGAGACGACCATGGCCA  180

Qy   181  ACCTGGCCGCGGTAAAACTGGGGTGAGCTGACCGCGCGGTGTGACGTAATGCCCTATCTC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  ACCTGGCCGCGGTAAAACTGGGGTGAGCTGACCGCGCGGTGTGACGTAATGCCCTATCTC  240

Qy   241  GGACATTACTGGCCGTCAGCGCGTCTGGGTAGCCCACGGAAATTGCTGCGGCAGAGGGCTGCCG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  GGACATTACCAACGCCAGCGCGTCTGGGTAGCCCACGGATTGCTGCGGCAGAGGGCTGCCG  300

Qy   301  CAGTCTGGTCAGGTCGCGGATTATTGCTGATGCTGATGCGACGAGCGTACCCGTAAAACAGGCGGAA
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  CAGTCTGGTCAGGTCGCGGATTATTGCTGATGCTGATGCGACGAGCGTACCCGTAAAACAGGCGGA  360

Qy   361  GCCTTCGCCGCGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGCCAGATACG
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  GCCTTCGCCGCGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGCCAGATACG  420

Qy   421  TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAAACTGGGCGTTTGCCAACTGGATAACGGCG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAAACTGGGCGTTTGCCAACTGGATAACGGCG  480

Qy   481  AACGTGACTGACGCGGATCCTCGAGAGGGCAGGAGGGTCAATGGCTGACTTTACCGGGGCCAT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  AACGTGACTGACGCGGATCCTCGAGAGGGCAGGAGGGTCAATGGCTGACTTTACCGGGGCCAT  540
```

```
Db   481  AACGTGACTGACCGATCCTCAGCGAGGGCGAGGAGGGTCAATTCGTTACTTTACCGGGCAT 540

Qy   541  TATCAAACGGGCGTTTCGGCAACTGGAACGGGTCGGTGCTTAATTTCCGCAATCAAACTTGTGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  CGGCAAACGGGCGTTTCGGCAACTGGAACGGGTCGGTGCTTAATTTCCGCAATCAAACTTGTGC 600

Qy   601  CTTAAACGTGAGAAACAGGAGCAGAAAGTCGTTCATTAACGGCATTACCATCGGAACTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  CTTAAACGTGAGAAACAGGAGCAGAAAGTCGTTCATTAACGGCATTACCATCGGAACTC 660

Qy   661  AAGGTGAGGCGCGCGACCTGTGTCTCCATTAACCGGTCGCGATTCAATCGTGACG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  AAGGTGAGGCGCGCGACCTGTGTCTCCATTAACCGGTCGCGATTCAATCGTGACG 720

Qy   721  GAGATATTTCTCCTGCAACAAGCACAGGGAATCCCGGAGGGGTGGGGAAGGATCACC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  GAGATATTCTCCTGCAACAAGCACAGGGAAATCCCGGAGGGGTGGGGAAGGATCACC 780

Qy   781  GATTCACACCAGTGGACACACCCTTGCTAAGTTTGCATAACGGCGCAAATTTGATTGGTTCTACAA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  GATTCACACCAGTGGAACACACCCTTGCTAAGTTTGCATAACGGCGCAAATTTTATTGGTACAA 840

Qy   841  CGCACGCGACGAGGTTGCCCGCGAGCCCCGTTATTGAGATTGATCAAGACAGCG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  CGCACGCGACGAGGTTGCCCGCGAGCCCCGTTATTGAGATTGATCAAGACAGCG 900

Qy   901  TTGACGCCCATCCACCGCCAAAAACAGGCGTATGGGTGGTTGACATTACCCACTTCAGTGCTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  TTGACGCCCCATCCACCGCCAAAAACAGGCGTATGGGTGTGACATTACCCACTTCAGTACTG 960

Qy   961  TTTATGCCCGGACACAGCATTACTAATCTCGGCATAAATCTCGGCGGAGCTCAACTGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  TTTATTGCCCGGACACGATACTAATCTCGGCAAATCTCGGCGGAGCTCAACTGG 1020

Qy  1021  ACGCTTCCCGGTCAGCCGGATAACACACGCGCCCAGAGTGGTGGTTGGTTTGAACGCTGG 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  ACGCTTCCCGGTCAGCCGGATAACACACGCGCCCAGAGTGGTGGTTGGTTTGAACGCTGG 1080

Qy  1081  CGTCGGCTAAGGCGATAACAGCCAGTGGATTCAGGTTTCGGCTCGTTCTCCAGACTTTACAG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  CGTCGGCTAAGGCGATAACAGCCAGTGGATTCAGGTTTCGGCTCGTTCTCCAGACTTTACAG 1140

Qy  1141  CAGATGCGTGATTAAAACGCCCGCTGTCATTAAATACCGCCCGGAGAGGTGAAAACTGACC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  CAGATGCGTGATTAAAACGCCCGCTGTCATTAAATACCGCCCGGAGAGGTGAAAACTGACC 1200

Qy  1201  CTGGCAGGATGTGAAGAGCGGAAATGCCCAGGCCATGTGTTCGTTCGGGCAGGTTTTACGCCA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  CTGGCAGGATGTGAAGAGCGGAAATGCCCAGGCCATGTGTTCGTTCGGGCAGGTTTTACGCCA 1260

Qy  1261  ATCGTGAATGAAGCACGCCATACCGGCCGTGCAGTTTGAGATCTCATC 1306
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  ATCGTGAATGAAGCACGCCATACCGGCCGTGCAGTTTGAGATCTCATC 1306
```

RESULT 14
AD050301

ID   AD050301 standard; DNA; 1323 BP.
XX
AC   AD050301;
XX
DT   29-JUL-2004 (first entry)
XX
DE   Kangaroo rat Escherichia coli phytase 872PH1 mutant DNA.
XX
KW   phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW   phytate; animal feed; fish feed; dough; baking; mutant; gene; ds.
OS   Escherichia coli.
XX
FH   Key   Location/Qualifiers
FH
FT   CDS   1. .1323
FT   /*tag= a

Query Match   96.7%; Score 1265; DB 12; Length 1323;
Best Local Similarity 98.0%;
Matches 1280; Conservative   0; Mismatches 26; Indels   0; Gaps   0;

```
Qy   1  ATGAAAGCGGATCTTAATCCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy       1 ATGAAAGCGGATCTTAATCCCATTTTTATCTTCTGATTCCGTTAAACCCGCAATCTGCA 60

Db      61 TTCGCTGCTGCTTCAAGAGTGAGCCGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120
Qy      61 TTCGCTGCTTCAGAGTGAGCCGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120

Db     121 GTGCGTGCTCCAACCACAAGGCCACGCAACTGATGCAGACGATGTCACCCGACGCATGGCCA 180
Qy     121 GTGCGTGCTCCAACCACAAGGCCACGCAACTGATGCAGACGATGTCACCCGACGCATGGCCA 180

Db     181 ACCTGGCCCGGTAAAAACTGGGTGAGCTGACACGGGTGGTGAGCTGAGTAAGTCGGCTATCTC 240
Qy     181 ACCTGGCCCGGTAAAAACTGGGTGGGTTGACACCGCGGTGGTGAGCTAATCGCTATCTC 240

Db     241 GGACATTACTGCCGTGGTCTCGGTTCAGGTTCGGTCTGGTAGCCGGACGGAATTGCTGCCCGGCCGG 300
Qy     241 GGACATTACCAACGCCAGCCGGTTCTGGTTAGCCGGACGGAATTGCTGCCCGG 300

Db     301 CAGTCTGGTCAGGTCGCGAACTGGGTCTGGTAGCCGGACGGAATTGCTGCCCGGCCGGAA 360
Qy     301 CAGTCTGGTCAGGTCGCGAACTGGGTCTGGTAGCCGGACGGAATTGCTGCCCGGCCGGAA 360

Db     361 GCCTTCGCGCCGCCGGCGCTGGCCACCTGACTGTGTCAATAACCGTTACATAACCCGTACAGATACG 420
Qy     361 GCCTTCGCGCCGCCGGCGCTGGCCACCTGACTGTGTCAATAACCGTTACATAACCCGTACAGATACG 420

Db     421 TCCAGTCCCGATCCGGTTATTTAATCCCTCTAAAAACTGGGTTGCCAACTGGATAACGCG 480
Qy     421 TCCAGTCCCGATCCGGTTATTTAATCCCTCTAAAAACTGGGTTGCCAACTGGATAACGCG 480

Db     481 AACGTGACTGACGCGATCCTCGAGAGGGCAGGAGGGGTCATTGCTGACTTTACCGGGCAT 540
Qy     481 AACGTGACTGACGCGATCCTCGAGAGGGCAGGAGGGGTCATTGCTGACTTTACCGGGCAT 540

Db     541 CGGCAAACGGCGGTTTCGCGGAACTGGGCTGCTTAATTTTCCGCAAATCAAAACTTGTGC 600
Qy     541 TATCAAAACGGCGGTTTCGCGGAACTGGGCTGCTTAATTTTCCGCAAATCAAAACTTGTGC 600

Db     601 CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC 660
Qy     601 CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC 660

Db     661 AAGGTGAGCGCCGACTGTGTCTTCATTAACCGGTGCGGTAAGCCTCGCCATCAAATGCTGACG 720
Qy     661 AAGGTGAGCGCCGACTGTGTCTTCATTAACCGGTGCGGTAAGCCTCGCCATCAAATGCTGACG 720

Db     721 GAGAGTATTTCTCCCTTGCAACAAGCACAGGGGAATGCCCGGACGAGCCGGGGTGGGGAGGGATCACC 780
Qy     721 GAGAGTATTTCTCCCTTGCAACAAGCACAGGGGAATGCCCGGACGAGCCGGGGTGGGGAGGGATCACC 780

Db     781 GATTCACACCAGTGGAACACCTTGCTAAGGTTTTGCATAAACGGCCCAATTTGATTTGCTACAA 840
Qy     781 GATTCACACCAGTGGAACACCTTGCTAAGGTTTTGCATAAACGGCCCAATTTGATTTGCTACAA 840

Db     841 CGCACGGCCACGAGGTTGCCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 900
Qy     841 CGCACGGCCACGAGGTTGCCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 900

Db     901 TTGACCCCCCATCCACCGCAAAAAACAGGGCGTATGGTGTGACATTACCCACTTCAGTTGCTG 960
Qy     901 TTGACCCCCCATCCACCGCAAAAAACAGGGCGTATGGTGTGACATTACCCACTTCAGTTGCTG 960

Db     961 TTTATTGCCGGACACGGATACTAAATCTCGGCGCGCGGCCACTGGAGCTCAACTGG 1020
Qy     961 TTTATTGCCGGACACGGATACTAAATCTCGGCGCGCGGCCACTGGAGCTCAACTGG 1020

Db    1021 ACGCTTCCCGGTCAGCCGGATAACACGCCCGGATAACTCGGGCTGGTGAACTGGGTTTGAACCGCTGG 1080
Qy    1021 ACGCTTCCCGGTCAGCCGGATAACACGCCCGGATAACTCGGGCTGGTGAACTGGGTTTGAACCGCTGG 1080

Db    1081 CGTCGCGCTAAGCGATAACAGCCAGTGGATTCAGGGTTTCGCTGGTGGTCTTCTTCCAGGACTTTACAG 1140
Qy    1081 CGTCGCGCTAAGCGATAACAGCCAGTGGATTCAGGGTTTCGCTGGTGGTCTTCTTCCAGGACTTTACAG 1140

Qy    1141 CAGATGCGTGATTAAAACGCCCGCTGTTCATTAAATACGCCGCCCCGGAGAGGTGAAAACTGACC 1200
Db    1141 CAGATGCGTGATTAAAACGCCCGCTGTTCATTAAATACGCCGCCCCGGAGAGGTGAAAACTGACC 1200

Qy    1201 CTGGCCAGGATGTGAAGAGCGGAAATGCGGCAGCCAGTGTTCGTTCGTTGGCAGGGTTTTACGCAA 1260
Db    1201 CTGGCCAGGATGTGAAGAGCGGAAATGCGGCAGCCAGTGTTCGTTCGTTGGCAGGGTTTTACGCAA 1260

Qy    1261 ATCGTGAATGAAGCACGCACCGGCGCTGCAGTTTGAGATCTCATC 1306
Db    1261 ATCGTGAATGAAGCACGCACCGGCGCTGCAGTTTGAGATCTCATC 1306
```

RESULT 15

AD050303
ID   AD050303 standard; DNA; 1323 BP.
XX
AC   AD050303;
XX
DT   29-JUL-2004 (first entry)
XX
DE   Escherichia coli W phytase 875PH2 mutant DNA.
XX
KW   Phytaee; bacteria; thermal tolerance; protease stability; foodstuff;
KW   phytate; animal feed; fish feed; dough; baking; mutant; gene; ds.
OS   Escherichia coli.
XX
FH   Key   Location/Qualifiers
FH
FT   CDS   1..1323
FT   /*tag= a
FT   /product= "Mutant phytase enzyme"
FT   unsure   214..216
FT   /*tag= b
FT   /note= "Encodes Arg"
FT   mutation   replace(478,G)
FT   /*tag= c
FT   mutation   replace(527,A)
FT   /*tag= d
FT   mutation   replace(893,T)
FT   /*tag= e
FT   mutation   replace(895,G)
FT   /*tag= f
XX
PN   US2004091968-A1.
XX
PD   13-MAY-2004.
XX
PF   20-JUN-2003; 2003US-00601319.
XX
PR   13-AUG-1997; 97US-00910798.
PR   01-MAR-1999; 99US-00252914.
PR   13-APR-1999; 99US-00291931.
PR   25-MAY-1999; 99US-00318528.
PR   25-MAY-2000; 2000US-00580515.
PR   24-MAY-2001; 2001US-00866379.
XX
PA   (SHOR/) SHORT J M.
PA   (KRET/) KRETZ K.
PA   (GRAY/) GRAY K A.
PA   (BART/) BARTON N R.
PA   (GARR/) GARRETT J B.
PA   (ODON/) O'DONOGHUE E.
PA   (MATH/) MATHER E J.
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI   Mather EJ;
XX
DR   WPI; 2004-374952/35.
DR   P-PSDB; AD050292, AD050304.
XX
PT   Producing phytase, involves providing nucleic acid derived from bacteria

PT   encoding polypeptide having phytase activity, and expressing nucleic acid
PT   in yeast.
XX
PS   Disclosure; Page; 74pp; English.
XX
CC   The invention relates to a method for producing phytase that involves
CC   providing a nucleic acid encoding phytase derived from a bacteria, and
CC   expressing the nucleic acid in a yeast under conditions that allow
CC   expression of the enzyme in the yeast. The invention also relates to
CC   modified phytase enzyme which has improved thermal tolerance and protease
CC   stability at low pH. The phytase enzyme can be used in foodstuffs to
CC   improve the feeding value of phytate rich ingredients, and in diet of
CC   numerous animals including mammals, fowls and fishes, commercially
CC   significant mammals such as pigs, goats, laboratory rodents, commercially
CC   significant avian species such as chicken, ducks, doves, parrot, etc.,
CC   commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC   etc., in dough making and baking, in dietary aids for animals. The method
CC   provides easy manufacture of the active ingredient loaded biocompatible
CC   composition, higher yields and loading efficiency. The phytase
CC   incorporated in to the dietary aid is safe for animals. Note: The phytase
CC   sequence is Escherichia coli W phytase mutant DNA. Note: The present
CC   sequence is not shown in the specification but has been derived from
CC   Escherichia coli B phytase DNA AD050291.
XX
SQ   Sequence 1323 BP; 324 A; 353 C; 356 G; 289 T; 0 U; 1 Other;

Query Match   96.7%;   Score 1265;   DB 12;   Length 1323;
Best Local Similarity   98.0%;   Pred. No. 0;
Matches 1280;   Conservative   0;   Mismatches   26;   Indels   0;   Gaps   0;

Qy   1 ATGAAAAGGCATCTTAATCCCATTTAATCTTCTGATTCCGTTAACCCGCAATCTGCA 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAAGGCATCTTAATCCCATTTAATCTTCTGATTCCGTTAACCCGCAATCTGCA 60

Qy   61 TTCGCTCAGAGTGAGCCGGAGTGAAGGTGGAAAAGTGTGGTGTCAGTGGT 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGTGGTTGAATTGTCAGTCGTCATCGGT 120

Qy   121 GTGCGTGCTCCAAACCAAGGCCACGCAACTGACTGAGTCAGGATGTCACCCAGACGGCATGGCCA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTGCGTGCTCCAAACCAAGGCCACGCAACTGACTGAGTCAGGATGTCACCCAGACGGCATGGCCA 180

Qy   181 ACCTGGCCGGGTAAAACTGGGTTGGCTGACACCCGCGNGGTGAGCTAATCGGCTACTC 240
            |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   181 ACCTGGCCGGGTAAAACTGGGTTGGCTGACACCCGCGNGGTGAGCTAATCGGCTACTC 240

Qy   241 GGACAATTACTGGCGTCAGCGTCTGGTAGCCGGACGAGGATTGCTCGCCTAAAATGTGGCCTGCCCG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   241 GGACAATTACCAAACGCCAGCGTCTGGTAGCCGGACGGATTGCTCGGCGAAAAAAGGCTGCCCG 300

Qy   301 CAGTCTGGTCAGGTCGGCGGATTATTGTGCGATGGTTCGACGAGCGTACCCCGTAAAAACAGGCGAA 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGTCTGGTCAGGTCGGCGGATTATTGTGCGATGGTTCGACGAGCGTACCCCGTAAAAACAGGCGAA 360

Qy   361 GCCTTCGCCGCGCGCGGCGGCGCTGGCACCTGACTGTGCAATAACCGTACAAACCCAGGCAGATACG 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCCTTCGCCGCGCGCGGCGGCGCTGGCACCTGACTGTGCAATAACCGTACAAACCCAGGCAGATACG 420

Qy   421 TCCAGTCCGATCCGGTTATTTTAATCCTTCTTAAAAATCTGGATAACGCG 480
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCCAGTCCGATCCGGTTATTTTAATCCTTCTTAAAAATCTGGATAACGCG 480

Qy   481 AACGTGACTGACGGCGGATCCTCAGCAGGGTCAAATTGCTCAACTTTTACCCGGGCAT 540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AACGTGACTGACGGCGGATCCTCAGCAGGGTCAAATTGCTCAACTTTTACCCGGGCAT 540

Qy   541 TATCAAACGGCGTTTCCCGAACTGGGAAGCGGTGTCGCTAATTTTCCGCAATCAAACTTGTGC 600
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CGGCAAACGGGCGTTTCCGGACTGGGAACTGGAACGGGGTCGCTAATTTTCGCAATCAAACTTGTGC 600

Qy   601 CTTAAAACGTGAGAAACAGGACGGAAAGCTGTTCATTAACGCAGGCCATTACCCATCGGAACTC 660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTTAAAACGTGAGAAACAGGACGGAAAGCTGTTCATTAACGCAGGCCATTACCCATCGGAACTC 660

Qy   661 AAGGTGAGCGCCCGACTGTGTCTCATTAACCGGTGTGCGGTAAGCCTCGCATCAAATGCTGACG 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AAGGTGAGCGCCCGACTGTGTCTCATTAACCGGTGTGCGGTAAGCCTCGCATCAAATGCTGACG 720

Qy   721 GAGATATTTTCTCCTGCAACAGCACAGGGAAATGCCGGACCCGGGTGGGGGAAGGATCACC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GAGATATTTTCTCCTGCAACAGCACAGGGAAATGCCGGACGGCGGGGTGGGGAAGGATCACC 780

Qy   781 GAATTCACCACCAGGTGGAAACACCTTGCTAAGTTTGCAATAACGGCGCAATTTGATTTGCTACAA 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GAATTCACCACCAGGTGGAAACACCTTGCTAAGTTTGCAATAACGGCGCAATTTTATTTGCTACAA 840

Qy   841 CGCACGCCAGAGGTTGCCCGCCAGCGGCCACCCGTTATTTGATTCAGACAGCGCGG 900
            ||||||||||||||||||||||||||||||||||||||||| ||| ||||||||||
Db   841 CGCACGCCAGAGGTTGCCCGGCAGCGGCCACCCGTTATTGGATTCAAGACAGCAGCG 900

Qy   901 TTGACGCCCCATCCACCCGCAAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCTG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTGACGCCCCATCCACCCGCAAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCTG 960

Qy   961 TTTATCGCCGGACACGGATACTAATCTGGCAAAATCTCGGGCGGGGCCGCACTCGGAGCTCAACTGG 1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TTTATTGCCGGACACGGATACTAATCTGGCAAAATCTCGGGCGGGGCCGCACTCGGAGCTCAACTGG 1020

Qy   1021 ACGCTTCCCGGTCAGCCGGATAACACGCCCGGCCAGGTGGTGGTTGGAACTGGTGTTTGAACGCTGG 1080
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGCTTCCCGGTCAGCCGGATAACACGCCCGGCCAGGTGGTGGTTGGAACTGGTGTTTGAACGCTGG 1080

Qy   1081 CGTCGGGCTAAGCGATAACAGCCCAGTCGGATTCAATTCAGGTTTCGGCTGGTTCTTCCAGACTTTACAG 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CGTCGGGCTAAGCGATAACAGCCCAGTCGGATTCAATTCAGGTTTCGGCTGGTTCTTCCAGACTTTACAG 1140

Qy   1141 CAGATGCGTGATAAAACCCGCTGTCATTAAATACGCCCCCCGGAGAGGTGAAAACTGACC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAGATGCGTGATAAAACCCGCTGTCATTAAATACGCCCCCCGGAGAGGTGAAAACTGACC 1200

Qy   1201 CTGGCAGGATGTGAAGAGACGGAAAATGCCGCAGGGCCATGGTGTTTCGTTCGGTTTGCAGGCAA 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGCAGGATGTGAAGAGACGGAAAATGCCGCAGGGCCATGGTGTTTCGTTCGGTTTTTACGGCAA 1260

Qy   1261 ATCGTGAATGAATGAAGCACGCGCCATTACCGGGCGTGCAGTTTGAGATCTCATC 1306
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAATGAAGCACGCGCCATTACCGGGCGTGCAGTTTGAGATCTCATC 1306

Search completed: June 14, 2006, 10:31:29
Job time : 893.897 secs

THIS PAGE LEFT BLANK

```
                                          GenCore version 5.1.9
                                   Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on:              June 14, 2006, 10:00:55 ; Search time 7316.05 Seconds
                                                     (without alignments)
                                              11432.839 Million cell updates/sec

Title:               US-10-601-319-9
Perfect score:       1308
Sequence:            1 atgaaagggatcctaatccc...............gcagttgagatctcaatta 1308
Scoring table:       IDENTITY_NUC
Searched:            6366136 seqs, 3197371.0525 residues
                         Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters:          12732272

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       GenEmbl:*
                 1: gb_env:*
                 2: gb_pat:*
                 3: gb_ph:*
                 4: gb_pl:*
                 5: gb_pri:*
                 6: gb_ro:*
                 7: gb_sts:*
                 8: gb_sy:*
                 9: gb_un:*
                10: gb_vi:*
                11: gb_ov:*
                12: gb_htg:*
                13: gb_in:*
                14: gb_om:*
                15: gb_ba:*
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1265	96.7	1323	2	AR108133	AR108133 Sequence
2	1265	96.7	1323	2	AR127818	AR127818 Sequence
3	1265	96.7	1323	2	AR130956	AR130956 Sequence
4	1265	96.7	1323	2	AR492966	AR492966 Sequence
5	1265	96.7	1323	2	AR636185	AR636185 Sequence
6	1265	96.7	1323	2	AX052713	AX052713 Sequence
7	1265	96.7	1323	2	AX338966	AX338966 Sequence
8	1265	96.7	1323	2	AX356566	AX356566 Sequence
9	1265	96.7	110000	15	U00096_10	Continuation (11 o
10	1264.8	96.7	110000	15	AP009048_10	Continuation (11 o
11	1263.8	96.6	1901	2	AR636189	AR636189 Sequence
12	1263.8	96.6	1901	2	AR636190	AR636190 Sequence
13	1263.8	96.6	1901	2	AX356571	AX356571 Sequence
14	1263.8	96.6	1901	2	AX356572	AX356572 Sequence
15	1263.8	96.6	1901	2	M58708	M58708 Escherichia
16	1263.2	96.6	110000	15	CP000038_10	Continuation (11 o
17	1262.2	96.5	1901	2	AR636188	AR636188 Sequence
18	1262.2	96.5	1901	2	AR636191	AR636191 Sequence
19	1262.2	96.5	1901	2	AX356570	AX356570 Sequence
20	1265	96.5	1901	2	AX356574	AX356574 Sequence
21	1262.2	96.5	1901	2	L03371	L03371 Escherichia
22	1261.6	96.5	1315	15	AF537219	AF537219 Escherich
23	1260.6	96.4	1901	2	L03372	L03372 Escherichia
24	1260.6	96.4	1901	2	L03373	L03373 Escherichia
25	1260.6	96.4	1901	2	L03374	L03374 Escherichia
26	1260.6	96.4	1901	2	L03375	L03375 Escherichia
27	1259.2	96.3	3470	2	AX042376	AX042376 Sequence
28	1259.2	96.3	4060	2	AX042375	AX042375 Sequence
29	1259.2	96.3	6116	2	AX042374	AX042374 Sequence
30	1259.2	96.3	6708	2	AX042373	AX042373 Sequence
31	1259.2	96.3	17732	2	AX042378	AX042378 Sequence
32	1259.2	96.3	20623	2	AX042372	AX042372 Sequence
33	1259	96.3	1901	2	L03370	L03370 Escherichia
34	1246.2	95.3	5421	2	AX042377	AX042377 Sequence
35	1239.2	95.2	110000	15	CP000036_22	Continuation (23 o
36	1239.2	94.7	110000	15	AE005674_10	Continuation (11 o
37	1237.6	94.6	110000	15	AE014073_10	Continuation (11 o
38	1233.6	94.3	1486	2	AR630587	AR630587 Sequence
39	1233.6	94.3	1489	2	AR277753	AR277753 Sequence
40	1228	93.9	1489	2	AY496073	AY496073 Escherich
41	1228	93.9	1486	15	AR630588	AR630588 Sequence
42	1222.4	93.5	1486	2	AE005174_12	Continuation (13 o
43	1218.8	93.2	110000	15	AE005174_13	Continuation (14 o
44	1218.8	93.2	110000	15	BA000007_12	Continuation (13 o
45	1218.8	93.2	110000	15	CP000034_08	Continuation (9 of
```
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

RESULT 1
AR108133

```
LOCUS       AR108133    1323 bp    DNA    linear    PAT 14-FEB-2001
DEFINITION  Sequence 1 from patent US 6110719.
ACCESSION   AR108133
VERSION     AR108133.1    GI:12823620
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1323)
  AUTHORS   Kretz,K.
  TITLE     Phytase
  JOURNAL   Patent: US 6110719-A 1 29-AUG-2000;
FEATURES             Location/Qualifiers
     source          1..1323
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN
```

```
Query Match             96.7%;  Score 1265;  DB 2;  Length 1323;
Best Local Similarity   98.0%;
Matches 1280;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;
```

```
Qy          1 ATGAAAGGGATCTTAATCCCATTTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAAGGGATCTTAATCCCATTTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60

Qy         61 TTCGCTCAGAGTGAGCCCGGAGCTGAAGCTGGAGAAAGTGTGGTTCAGTTCGTTCATCGGT 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGCTCAGAGTGAGCCCGGAGCTGAAGCTGGAAAGTGTGGTTGTTCAGTTCGTTCATCGGT 120

Qy        121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGAGTCACCCCAGACGGCATGGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGAGTCACCCCAGACGGCATGGCCA 180

Qy        181 ACCTTGGGCCGGTAAAACTGGGTGAGCTGACACCGCGCGTAATCCGCCTATCTC 240
              |||||||||||||| |||||||||||||||||||||||||||||||| ||||
Db        181 ACCTTGGGCCGGGTTAAAACTGGGTTGGCTGACACCGCGNGGCTAATCCGCCTATCTC 240

Qy        241 GGACAATTACTGGCGGTCAGCGTCTCGGTAGCGCGACGGATTGCTGCCTAAATGTGGCGCCG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ...
```

ALIGNMENTS

RESULT 2

```
LOCUS       AR127818                1323 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 1 from patent US 6183740.
ACCESSION   AR127818
VERSION     AR127818.1  GI:14115480
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1323)
  AUTHORS   Short,J.M. and Kretz,K.A.
  TITLE     Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: US 6183740-A 1 06-FEB-2001;
            Recombinant bacterial phytases and uses thereof
FEATURES             Location/Qualifiers
     source          1..1323
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Query Match            96.7%;  Score 1265;  DB 2;  Length 1323;
Best Local Similarity  98.0%;  Pred. No. 0;
Matches 1280;  Conservative   0;  Mismatches  26;  Indels   0;  Gaps   0;
```

```
Qy    1 ATGAAACGATCTTAATCCCATTTTAATCCTTCTTCTGATTCGTTAACCCCGCAAATCTGCA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAACGATCTTAATCCCATTTTAATCCTTCTTCTGATTCGTTAACCCCGCAAATCTGCA 60

Qy   61 TTCGCTCAGAGTGAGCCCGGAGCCTGAAGCTGAAAGTGTGGTGATTGCTGAGTTCGTCATGGT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCCGGAGCCTGAAGCTGAAAGTGTGGTGATTGCTGAGTTCAGTTGGT 120

Qy  121 GTGCGTGCTCCAACCAAGGCCACGCAAACTGAATGCAGGATGTCACCCCAGAGCGCATGGCCA 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTGCGTGCTCCAACCAAGGCCACGCAAACTGAATGCAGGATGTCACCCCAGAGCGCATGGCCA 180

Qy  181 ACCTGGCCGGTAAAACTGGGGTTGGCCTGACCGTGTGCAATAACCCAGGCAGATACG 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ACCTGGCCGGTAAAACTGGGGTTGGCCTGACCGTGTGCAATAACCCAGGCAGATACG 240

Qy  241 GGACATTACTGGCGTCAGCGTCTGGTAGCCACGGATTGCTGCCCCGG 300
        |||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GGACATTACTGGCGTCAGCGTCTGGTAGCCACGGATTGCTGCCCCGG 300

Qy  301 CAGTCTTGGTCAGGTCGCCGGATTATTGCTGATGTCGACGACGGAGCGGATTCGCCCGG 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGTCTTGGTCAGGTCGCCGGATTATTGCTGATGTCGACGACGGAGCGGATTCGCCCGG 360

Qy  361 GCCTTCGCCGCCGCGGGGCTGGGCACCTGACTGTTCAATAACCGTACATACCAGGCAGATACG 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GCCTTCGCCGCCGCGGGGCTGGGCACCTGACTGTTCAATAACCGTACATACCAGGCAGATACG 420

Qy  421 TCCAGTCCCGATCCGGTTATTTAATCCTCTAAAAAACTGGCCTGACGCG 480
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TCCAGTCCCGATCCGGTTATTTAATCCTCTAAAAAACTGGCCTGACGCG 480

Qy  481 AACGTGACTGACGCGGATCCTCAGCAGGGGTCAATTGCTGACTTTACCGGGCAT 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 AACGTGACTGACGCGGATCCTCAGCAGGGGTCAATTGCTGACTTTACCGGGCAT 540

Qy  541 TATCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCCGCAATCAAACTTGTGC 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCCGCAATCAAACTTGTGC 600

Qy  601 CTTAAACGTGAGAAACAGGACGGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAAACTC 660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CTTAAACGTGAGAAACAGGACGGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAAACTC 660

Qy  661 AAGGTGAGCGCCGACTTGTGTCTCATTAACCGGTGGTAAGCCTCGGCATCAATGCTGACG 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 AAGGTGAGCGCCGACTTGTGTCTCATTAACCGGTGGTAAGCCTCGGCATCAATGGTGACG 720

Qy  721 GAGATATTTCTTCCTGCAACAAGCACAGGAAGCCGGAGCGCGGGGGTGGGGAAGGAGATCACC 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GAGATATTTCTTCCTGCAACAAGCACAGGAAGCCGGAGCGCGGGGGTGGGGAAGGAGATCACC 780
```

```
Db  241 GGACATTACCAACGGCCAGCGGATTATTGCTGGTGAGTTCGTCATGGT 300
Qy  241 CGCAGGCCAGAGGTTGCCCCGACCGCGCCAGCCGCGCGGTAAGCTGGG 300

Db  301 CAGTCTTGGTCAGGTCGCCGGATTATTGCTGATGTCGACGACGGAGCGGA 360
Qy  301 CGTCCGCGGCTAAGCGGATACCAGCAGTGGATTCAGGGTTCTTCCGAGACTTTACAG 360

Db  361 GCCTTCGCCGCCGCGGGGCTGGGCAC 420
Qy  361 TTTATGCGGCGGACAGCTAATCTCGGGCTAAATCTCGGGAGCTCAACTGG 420

Db  421 TCCAGTCCCGATCCGGTT 480
Qy  421 ACGCTTCCCGGTCAGCGCGGA 480

Db  481 AACGTGACTGACGCGGATCCT 540
Qy  481 AACGTGACTGACGCGGATCCT 540

Db  541 CGGCAAACGGCGTTTCGCGAACTGGAA 600
Qy  541 CAGATACGGTTGATAAAACGCGCGGTGTCA 600

Db  601 CTTAAACGTGAGAAACAGGAC 660
Qy  601 CTTAAACGTGAGAAACAGGAC 660

Db  661 AAGGTGAGCGCCGAC 720
Qy  661 AAGGTGAGCGCCGAC 720

Db  721 GAGATATTTCTTCCTGCAACAAGCACAGGAAGCCGGAGCGCGGGGGTGGGGAAGGAGATCACC 780
Qy  721 GAGATATTTCTTCCTGCAACAAGCACAGGAAGCCGGAGCGCGGGGGTGGGGAAGGAGATCACC 780
```

RESULT 3

```
LOCUS       AR130956                1323 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 1 from patent US 6190897.
ACCESSION   AR130956
VERSION     AR130956.1  GI:14119281
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1323)
  AUTHORS   Kretz,K.
  TITLE     Phytase
  JOURNAL   Patent: US 6190897-A 1 20-FEB-2001;
FEATURES             Location/Qualifiers
     source          1..1323
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN
```

```
Query Match             96.7%;  Score 1265;  DB 2;  Length 1323;
Best Local Similarity   98.0%;
Matches 1280;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;
```

```
Qy          1 ATGAAAGGCGATCTTAATCCCATTTTATCTCTTCCGTTACCCGCCCAAATCTGCA     60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAAGGCGATCTTAATCCCATTTTATCTCTTCCGTTACCCGCCCAAATCTGCA     60

Qy         61 TTCGCTCAGAGTGAGCTGGAAAGTGTGGATTGGTGAGTGTCAGTCGTGCTCATGGT    120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGCTCAGAGTGAGCTGGAAAGTGTGGATTGGTGAGTGTCAGTCGTGCTCATGGT    120

Qy        121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGAGTGTCACCCGACGCGATGGCCA   180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGAGTGTCACCCGACGCGATGGCCA   180
```

(alignment continues in paired Qy/Db blocks through position 1306/1260)

```
RESULT 4
AR492966

LOCUS       AR492966     1323 bp    DNA     linear   PAT 15-MAY-2004
DEFINITION  Sequence 1 from patent US 6720014.
ACCESSION   AR492966
VERSION     AR492966.1  GI:47264255
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1323)
  AUTHORS   Short,J.M. and Kretz,K.A.
  TITLE     Phytase-containing foodstuffs and methods of making and using them
  JOURNAL   Patent: US 6720014-A 1 13-APR-2004;
            Diversa Corporation; San Diego, CA
FEATURES             Location/Qualifiers
     source          1..1323
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             96.7%;  Score 1265;  DB 2;  Length 1323;
  Best Local Similarity   98.0%;  Pred. No. 0;  Mismatches 26;  Indels 0;  Gaps 0;
  Matches 1280;  Conservative   0;

Qy      1 ATGAAAGCGGATCTTAATCCCATTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAGCGGATCTTAATCCCATTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60

Qy     61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGGGTGATTGTCAGTCGTCATGGGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGTGGTGATTGTCAGTCGTCATGGGT 120

Qy    121 GTGCGTGCTCCAACCAAGGGCCACGCAAACTCATGGGATCAGGTGACGCATGGCCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGCGTGCTCCAACCAAGGGCCACGCAAACTCATGGGATCAGGTGACGCATGGCCA 180

Qy    181 ACCTTGGCCCGGTAAAAACTGGGTTGGCTGACACCCGCCGCCTAATCGCCTATCTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACCTTGGCCCGGTAAAAACTGGGTTGGCTGACACCCGCCGCCTAATCGCCTATCTC 240

Qy    241 GGACATTACTGCGTCAGCGGTCTGGTAGCCGACGGATTGCTGCCTAAAGTGGCTGCCCG 300
          ||||||||||||||||||||||| | ||||||||||||||||||||||||||||||||
Db    241 GGACATTACCAAACGCCAGCGTCTCGTAGCCGACCGGATTGCTGGCGAAAAAGGGCTGCCCG 300

Qy    301 CAGTCTGGTCAGGTCGGCGAGTATTGCTGCTGACGAGCGGTACCCGTAAAACGACGGAA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGTCTGGTCAGGTCGGCGAGTATTGCTGCTGACGAGCGGTACCCGTAAAACGACGGAA 360

Qy    361 GCCTTCGCCCGCCGGCGGCTGGGCACCTGACTGTGCAATAACCGTACATACCAGGGCATACG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCTTCGCCCGCCGGCGGCTGGGCACCTGACTGTGCAATAACCGTACATACCAGGCGCATACG 420

Qy    421 TCCAGTCCCCGGATCCGTTATTTAATCCTTCTAAAAAACTGGGATAACGGCG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCCAGTCCCCGGATCCGTTATTTAATCCTTCTAAAAAACTGGGATAACGGCG 480

Qy    481 AACGTGACTGACGCGCGATCCTCAGCAGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AACGTGACTGACGCGCGATCCTCAGCAGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 540

Qy    541 TATCAAAACGGCGTTTCGCGGAACTGGAACGGGCTGGTTCATTAATTTTCCCGAATCAAAACTTGTGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CGGCAAAACGGCGGTTCGCGGAACTGGAACGGGCTGGTTCATTAATTTTCCCGAATCAAAACTTGTGC 600

Qy    601 CTTAAACGTGAGAAACAGGAGGACGAAAAAGCTGTTCATTAACGGAGGCATTACGTCGGAAACTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTTAAACGTGAGAAACAGGAGGACGAAAAAGCTGTTCATTAACGGAGGCATTACGTCGGAAACTC 660
```
```
Qy   1261 ATCGTGAATGAAGCACGCATACCGGCGGTGCAGTTTGAGATCTCATC 1306
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAGCACGCATACCGGCGGTGCAGTTTGAGATCTCATC 1306
```

```
Qy    661 AAGGTGAGCGGCCGACTGTGTCTCATTAAACCGGTGCCGTAAGCCTCGGCATCAAATGCTGACG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AAGGTGAGCGGCCGACTGTGTCTCATTAAACCGGTGCCGTAAGCCTCGGCATCAAATGCTGACG 720

Qy    721 GAGATATTTCTCCTGCAACAAGCACAGAGGAATGCCCGAGCGAGCCGGAGGGATCACC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCCGAGCGAGCCGGAGGGATCACC 780

Qy    781 GATTCACACCAGTGGAACACCTTGCTAAGTTGCATAAACGGCAATTGATTTGCTACAA 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GATTCACACCAGTGGAACACCTTGCTAAGTTGCATAAACGGCAATTGATTTGCTACAA 840

Qy    841 CGCACCGCCAGAGGTTGCCCGCGGCAACCCCGTTATTAGATTTGATCAAGACAGCG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CGCACCGCCAGAGGTTGCCCGCGGCAACCCCGTTATTAGATTTGATCAAGACAGCG 900

Qy    901 TTGACGCCCCATCCACCCGCAAAAAACAGGGCGTATGGGTTGACATTACCCACTTACTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTGACGCCCCATCCACCCGCAAAAAACAGGGCGTATGGGTTGACATTACCCACTTACTG 960

Qy    961 TTTATCGGCCGGACACGGATACTAATCTGGCAAAATCTCGGCGTGACGCACTGGAACGCTGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTTATCGGCCGGACACGGATACTAATCTGGCAAAATCTCGGCGTGACGCACTGGAACGCTGG 1020

Qy   1021 ACGCTTCCCGGTCAGCGGATAAACACGCCGGGTGGTGAACTGGTGTTGTTTGAAACGGCTGG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGCTTCCCGGTCAGCGGATAAACACGCCGGGTGGTGAACTGGTGTTGTTTGAAACGCTGG 1080

Qy   1081 CGTCCGGCCTAAGCGCATAAACAGCCAGTGGATTCAGGTTTCAGCTTCAGACTTTACAG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CGTCCGGCCTAAGCGCATAAACAGCCAGTGGATTCAGGTTTCAGCTTCAGACTTTACAG 1140

Qy   1141 CAGATCCGGTGATAAAACGCCCGCCTGTCATTAAAATAACGGCCGCCCGGAGAGGTGAAAACTGACC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAGATCCGGTGATAAAACGCCCGCCTGTCATTAAAATAACGGCCGCCCGGAGAGGTGAAAACTGACC 1200

Qy   1201 CTGGCAGGATGTGAAGAGCGGAAATGGGCCAGGGCCATGTGGTTCGTTGGCAGGTTTTACGGCAA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGCAGGATGTGAAGAGCGGAAATGGGCCAGGGCCATGTGGTTCGTTGGCAGGTTTTACGGCAA 1260
```

```
RESULT 5
AR636185

LOCUS       AR636185     1323 bp    DNA     linear   PAT 20-APR-2005
DEFINITION  Sequence 1 from patent US 6855365.
ACCESSION   AR636185
VERSION     AR636185.1  GI:62768868
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1323)
  AUTHORS   Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B.,
            O'Donoghue,E. and Mathur,E.J.
  TITLE     Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: US 6855365-A 1 15-FEB-2005;
            Diversa Corporation; San Diego, CA
FEATURES             Location/Qualifiers
     source          1..1323
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             96.7%;  Score 1265;  DB 2;  Length 1323;
  Best Local Similarity   98.0%;  Pred. No. 0;  Mismatches 26;  Indels 0;  Gaps 0;
  Matches 1280;  Conservative   0;
```

```
Qy      1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTTGATTCCGTTAACCCCGCAATCTGCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTTGATTCCGTTAACCCCGCAATCTGCA 60

Qy     61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120

Qy    121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy    181 ACCTGGCCGGGTAAAACTGGGTGGCTGAGCTGACACACCCGAGAGGCTAATCGGCCTATCTC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACCTGGCCGGGTAAAACTGGGTGGCTGAGCTGACACACCCGAGAGGCTAATCGGCCTATCTC 240

Qy    241 GGACATTACTACGGCGTCAGGTGGCTCAGGCGTTCAGGTAGCCGACGGAATTGCTGCCCG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGACATTACCAACGGCGTCAGGTGGCTCAGGCGTTCAGGTAGCCGACGGAATTGCTGCCCG 300

Qy    301 CAGTCTGGTCAGGTCGCGGATTATTGCTGATGTCGACGACGGGTACCCGTTAAAACAGG 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGTCTGGTCAGGTCGCGGATTATTGCTGATGTCGACGACGGGTACCCGTTAAAACAGG 360

Qy    361 GCCTTCGCGCGCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCTTCGCGCGCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420

Qy    421 TCCAGTCCGATCCGTTATTTAATCCTCTAAAAACTGGGCGTTGCCAACTGGATAAGCGG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCCAGTCCGATCCGTTATTTAATCTCTAAAAACTGGGCGTTGCCAACTGGATAAGCGG 480

Qy    481 AACGTGACTGACCGGATCCTCAGAGGGGCAGAGGGGTCAATTGCTTACCGGGGCAT 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AACGTGACTGACCGGATCCTCAGAGGGGCAGAGGGGTCAATTGCTTACCGGGGCAT 540

Qy    541 TATCAAACGGCGGTTTCGGGAACTGGGAACGGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CCGCAAACGGGCGGTTTCGGGAACTGGGAACGGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600

Qy    601 CTTAAACGTGAGAAAACAAGGACGAAAGCTGTTCATTAAGCGGCAGCCATTACCATGGAACTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTTAAACGTGAGAAAACAAGGACGAAAGCTGTTCATTAAGCGGCAGCCATTACCATGGAACTC 660

Qy    661 AAGGTGAGCGCCCACTGTGTCTCATTAACCGGTGGCGGTAAGCCTTCGCCGCATCAATGCTGACG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AAGGTGAGCGCCCACAATGTCTCATTAACCGGTGGCGGTAAGCCTTCGCCGCATCAATGCTGACG 720

Qy    721 GAGATATTTCTCCTGCAACAGCACAGGGAATGCCGGAGACCCGGGTGGGGAAGGATCACC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GAGATATTTCTCCTGCAACAGCACAGGGAATGCCGGAGACCCGGGTGGGGAAGGATCACC 780

Qy    781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAAAGCGCGCAATTTGATTGTACTACA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAAAGCGCGCAATTTATTGTACTACA 840

Qy    841 CGCACGCCAGAGGTTGCCCGCGCCAGCCCGCCACCCGGTTGATTTGATCAAGACAGCG 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CGCACGCCAGAGGTTGCCCGCGCCAGCCCGCCACCCGGTTATTTGATCAAGACAGCG 900

Qy    901 TTGACCGGCCCATCCACCGCAAAAACAGGGCGTATGGTGTGACATTACCCACTTCAGTGCTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTGACCGGCCCATCCACCGCAAAAACAGGGCGTATGGTGTGACATTACCCACTTCAGTACTG 960

Qy    961 TTTATTGCCCGGACAACGATACTAATCTCGGCGAAATCTCGCGGCCACTGGACTCAACTG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTTATCCGGACAACGATACTAATCTCGGCGAAATCTCGCGGCCACTGGACTCAACTG 1020

Qy   1021 ACGCTTCCGGTCAGCCGGATAACACCGCCCAGGTGGTGAACTGGTGTTTGAACGCGTGG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGCTTCCGGTCAGCCGGATAACACCGCCCAGGTGGTGAACTGGTGTTTGAACGCGTGG 1080

Qy   1081 CGTCGGCTTAAGCCGATAACAGCCAGTGGAATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CGTCGGCTTAAGCCGATAACAGCCAGTGGAATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140

Qy   1141 CAGATGCGTGATAAAAACGCCGCGTTCATTCATTAAATATACGGCGTTTCGCTGGTCTTCCAGACTTTACAG 1200
          ||||||||||||||||||||||||||||||||||||||||
Db   1141 CAGATGCGTGATAAAAACGCCGCGTTCATTCATTAAATATACGGCGTTTCGCTGGTGAAAACTGACC 1200

Qy   1201 CTGGCCAGGATGTGAAGAGAGCCGGAAATGCGCACAGGGCCATGTGTTCGTTGGCAGTTTACGGCAA 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGCCAGGATGTGAAGAGAGCCGGAAATGCGCACAGGGCCATGTGTTCGTTGGCAGTTTACGGCAA 1260

Qy   1261 ATCGTGAATGAAGCACGCATACCCGGCCGTGCAGTTTGAGATCTCATC 1306
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAGCACGCATACCCGGCCGTGCAGTTTGAGATCTCATC 1306
```

RESULT 6

```
LOCUS       AX052713                1323 bp    DNA     linear   PAT 12-JAN-2001
DEFINITION  Sequence 1 from Patent WO0071728.
ACCESSION   AX052713
VERSION     AX052713.1  GI:12226903
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Short,J.M. and Kretz,K.A.
  TITLE     Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: WO 0071728-A 1 30-NOV-2000;
            Diversa Corporation (US)
FEATURES             Location/Qualifiers
     source          1..1323
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
     CDS             1..1323
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC21810.1"
                     /db_xref="GI:12226904"
                     /translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP
                     TKATQLMQDVTPDAWPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQS
                     GQVAIIADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNA
                     NVTDAILSRAGGSIADPTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPS
                     ELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHOWNTLLSLHNAQF
                     YLLQRPEVARSRATPLLDLIMAALTPHPPQKQAYGVTLPTSVLFIFAGHDTNLANLGG
                     ALELNMTLPGQPDNTPGGELVFERRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNT
                     PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSLRSHHHHHH"
ORIGIN

Query Match            96.7%;   Score 1265;  DB 2;  Length 1323;
Best Local Similarity  98.0%;   Pred. No. 0;  Mismatches 26;
Matches 1280;  Conservative 0;  Mismatches 26;  Indels 0;  Gaps 0;
```

```
Qy      1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTTGATTCCGTTAACCCCGCAATCTGCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTTGATTCCGTTAACCCCGCAATCTGCA 60

Qy     61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120

Qy    121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy    181 ACCTGGCCGGGTAAAACTGGGTTGGCTGAGCTGACACACCCGNGGTTGGTGAGCTAATCGGCCTATCTC 240
          |||||||||||||||||||||||||||||||||||||
Db    181 ACCTGGCCGGGTAAAACTGGGTTGGCTGAGCTGACACACCCGNGGTTGGTGAGCTAATCGGCCTATCTC 240.
```

```
RESULT 7
LOCUS       AX338966               1323 bp    DNA        linear   PAT 09-JAN-2002
DEFINITION  Sequence 1 from Patent WO0189317.
ACCESSION   AX338966
VERSION     AX338966.1  GI:18129102
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Short,J.M., Kretz,K.A. and O'Donoghue,E.
  TITLE     Dietary aids and methods of use thereof
  JOURNAL   Patent: WO 0189317-A 1 29-NOV-2001;
            DIVERSA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..1323
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
     CDS             1..1323
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /db_xref="GI:18129103"
                     /protein_id="CAD20553.1"
                     /translation="MKAIIIPFLSLLIPITPQSAFPAQSEPELKLESVVIVSRHGVRAP
                     TKATQLMQDVTPDAWPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQS
                     GQVAIIADVDERTRKTGEAPAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA
                     NVTDAILSRAGGSIADFTGHRQTAPRELERVLNFPQSNCLKREKQDESCSLTQALPS
                     ELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQF
                     YLLQRTPEVARSRATPLLDLIMAALTIHPPQKQAYGVTLPTSVLFIAGHDTNLANLGG
                     ALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSIVFQTLQQMRDKTPLSLNT
                     PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSLRSHHHHHH"
ORIGIN
Query Match     96.7%;  Score 1265;  DB 2;  Length 1323;
Best Local Similarity   98.0%;
Matches 1280;  Conservative   0;  Mismatches   26;  Indels   0;  Gaps   0;

Qy      1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCCAANCTGCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db      1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCCAATCTGCA 60

Qy     61 TTCGCTCAGAGTGAGCCGGCGGAGCTGAAAAGTGTGGTGATTGTCAGTCGTCATGGT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db     61 TTCGCTCAGAGTGAGCCGGCGGAGCTGAAAAGTGTGGTGATTGTCAGTCGTCATGGT 120

Qy    121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||| 
Db    121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTGACCCCAGACGCATGGCCA 180

Qy    181 ACCTGGCCGGTAAAAACTGGGGTGAGCTGACACCGCGCGGGGGTGGTGAGCTGATACGCTC 240
          ||||||||||||||||||||||||||||||||||||||||||| |||||||||||| |||| 
Db    181 ACCTGGCCGGTAAAAACTGGGGTGAGCTGACACCGCGCGGGGGNGGTGAGCTGATATCTC 240

Qy    241 GGACATTACTTGGCGTCAGCGTCTGGTAGCCGGACGGATTGCTGCCTAAATGTGGCTGCCCG 300
          |||||||| ||| || ||||| ||| |||||||||| |||||| || |||   |||||||| 
Db    241 GGACATTACCAACGCCAGCGTTCGGTTAGCCGGACCGACGGAATTGCTGCCGGAAAAAGGGCTGCCCG 300

Qy    301 CAGTCTGGTCAGGTCGCCGGATTATTGCTGATGTCGACGAGGCCTACCCGTAAAACAGCGGAA 360
          |||||||||||||| | ||||||||||||| |||||| ||||||||||| |  ||||||||| 
Db    301 CAGTCTGGTCAGGTCTCCGGATTATTGCTGATGTCGACGAGGCCTACCCGTAAAACAGCGGAA 360

Qy    361 GCCTTCGCCGCGCCGGGGCTGGCACCTGACTGTGCGAATAACCGTACATACCGAGGCAGATACG 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    361 GCCTTCGCCGCGCCGGGGCTGGCACCTGACTGTGCAATAACCGTACATACCGAGGCAGATACG 420

Qy    421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAAACTGGATAACGCG 480
          ||||||| ||||||| |||| |||||| |||   ||||||||||||| 
Db    421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAAACTGGATAACGCG 480
```

```
Qy  241 GGACATTACTTGGCGTCAGCGTCTGGTAGCCGGACGGATTGCTGCCTAAATGTGGCTGCCCG 300
        |||||||| ||| || ||||| ||| |||||||||| |||||| || |||   |||||||| 
Db  241 GGACATTACCAACGCCAGCGTTCGGTTAGCCGGACCGACGGAATTGCTGCCGGAAAAAGGGCTGCCCG 300

Qy  301 CAGTCTGGTCAGGTCGCCGGATTATTGCTGATGTCGACGAGGCCTACCCGTAAAACAGCGGAA 360
        |||||||||||||| | ||||||||||||| |||||| ||||||||||| |  ||||||||| 
Db  301 CAGTCTGGTCAGGTCTCCGGATTATTGCTGATGTCGACGAGGCCTACCCGTAAAACAGCGGAA 360

Qy  361 GCCTTCGCCGCGCCGGGGCTGGCACCTGACTGTGCGAATAACCGTACATACCGAGGCAGATACG 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db  361 GCCTTCGCCGCGCCGGGGCTGGCACCTGACTGTGCAATAACCGTACATACCGAGGCAGATACG 420

Qy  421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAAACTGGATAACGCG 480
        ||||||| ||||||| |||| |||||| |||   ||||||||||||| 
Db  421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAAACTGGATAACGCG 480

Qy  481 AACGTGACTGACGCGGATCCTCTGAGAGTGGGGTCAATTGCTGATTTTACCGGGCAT 540
        |||||||||| |||| || ||| |||||| |||||| |||||||| ||||||||||| 
Db  481 AACGTGACTGACGCGGATCCTCAGCACGGGTCAATTGCTGATTTTACCGGGCAT 540

Qy  541 TATCAAACGGCCTTTCGGCGAACTGGAACGGGTGTTGAACTGGATAACCATCGAATTGTGC 600
        ||  |||||||| |||| |||||||||||||| ||||||||| ||||||| ||| 
Db  541 CGGCAAAACGGCGGTTTTCGGCAGAACTGGATTAATTTTCCGCAAAGCAATTGTGC 600

Qy  601 CTTAAACGTGAGAAACAGGACGAGAAAGCTGTTCATTAACGGCAAGCACTGGAAACTC 660
        ||||||||||||||||||| ||||||| |||||||| ||||| ||||| ||||||||| 
Db  601 CTTAAACGTGAGAAACAGGACGAGAAAGCTGTTCATTAACCGGTAAGCCTCGCATCAATGTCACG 660

Qy  661 AAGGTGAGCGCGCCGGACTGTGTGCTTCATTAACCGGTGCGGTAAGCCTCGCATCAATGTCACG 720
        |||||||||||||||| |||| ||||||||||||| ||||||||||||||||||||||||||| 
Db  661 AAGGTGAGCGCGCCGGACGCAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGTCACG 720

Qy  721 GAGATATTTCCTCGCAACAAGCACAGGAATGCCGGAGCGGGTGGGGAAGGAAGAGATCACC 780
        ||||||||| ||| ||||| ||||||| ||||||||||||||||||||||||||||||||| 
Db  721 GAGATATTTCCTGCAACAAGCCCGAGCGGGTGGGGAAGGAAGAGATCACC 780

Qy  781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATTAAACGGCGCAATTTATTTGCTACAA 840
        ||||||||||||||||||||||| |||||||||| ||||||| |||| |||| ||||||| 
Db  781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATTAAACGGCGCAATTTATTTGCTACAA 840

Qy  841 CGCACGCCAGAGGTTGCCCCGCAGCCCGCCGCCAATTTTATTAGAGATTTGATCATGGGCAGCG 900
        |||||||| ||||| ||||| |||||||||||||||||| ||||||| ||||||||||||||| 
Db  841 CGCACGCCAGAGGTTGCCCCGCAGCCCGCCGCCAATTTTATTAGAGATTTGATCATGGGCAGCG 900

Qy  901 TTGACCCCCATCCACCGCAAAAAACAGGCGTATGGTGTGACATTACCACTTCAGTGCTG 960
        ||||| ||||| |||| |||||||||| |||||| ||||||||||||||||||||||| 
Db  901 TTGACCCCCATCCACCGCAAAAAACAGGCGTATGGTGTGACATTACCACTTCAGTGCTG 960

Qy  961 TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGCGCTGGAGCTGAACTGG 1020
        |||| ||||||||||||| |||||||||||| ||||| |||| |||||||||||||| 
Db  961 TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGCGCTGGAGCTGAACTGG 1020

Qy 1021 ACGCTTCCCGGTCAGCCGGATAACACGCCGCCCGGAGGTGAACTGGTGTTTGAACGCTGG 1080
        ||||||||||||||||| |||||||||||| ||||||||| |||||||| ||||||| 
Db 1021 ACGCTTCCCGGTCAGCCGGATAACACGCCGCCCGGAGGTGAACTGGTGTTTGAACGCTGG 1080

Qy 1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTGTTTCAGACTTTACAG 1140
        ||||| |||||||||||||||||||||||||||| ||||||||| |||||||||||||| 
Db 1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTGTTTCAGACTTTACAG 1140

Qy 1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACCCCGCCGGGAGAGGTGAAAACTGACC 1200
        |||||||||||||||||||||||||||| ||||| ||||||||||| |||||||||| 
Db 1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACCCCGCCGGGAGAGGTGAAAACTGACC 1200

Qy 1201 CTGGCAGGATGTGAACAGCGAAATGGCACGGCCATGTGTTCGTTGGCAGGTTTTACGCAA 1260
        |||||||| |||||||||||| ||| |||||||||||||| |||||| ||||||||||| 
Db 1201 CTGGCAGGATGTGAACAGCGAAATGGCACGGCCATGTGTTCGTTGGCAGGTTTTACGCAA 1260

Qy 1261 ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTGAGATCTCATC 1306
        |||||||||||||||| |||||||||||||||||||||||||||| 
Db 1261 ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTGAGATCTCATC 1306
```

RESULT 8

```
LOCUS       AX356566               1323 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 1 from Patent WO0190333.
ACCESSION   AX356566
VERSION     AX356566.1  GI:18621053
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B. and
            O'Donoghue,E.
  TITLE     Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: WO 019033-A 1 29-NOV-2001;
```

```
FEATURES             Location/Qualifiers
     source          1..1323
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
     CDS             1..1323
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAD23025.1"
                     /db_xref="GI:18621054"
                     /translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP
                     TKATQLMQDVTPDAWPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQS
                     GQVAIIADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA
                     NVTDAILSRAGGSIADFTGHRQTAPRELKSRVLNFPQSNLCLKREKQDESCSLTQALPS
                     ELKVSADNVSITGAVSLASMITEIFLLQQAQGMPEPGWGRRITDSHQWNTLLSLHNAQP
                     YLLQRTPEVARSRATPLLDLIMAALTPHPPQKQAYGVTLPTSVLFIAGHDTNIANLGG
                     ALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNT
                     PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSLRSHHHHHH"
```

ORIGIN

```
Qy   Query Match   96.7%;  Score 1265;  DB 2;  Length 1323;
     Best Local Similarity   98.0%;
Matches 1280;  Conservative    0;  Mismatches   26;  Indels     0;  Gaps     0;
```

DIVERSA CORPORATION (US)

Qy   721 GAGATATTTCTTCCTTGCAACAAGCACAGGGAGAATGCCGGAGCCCGGGGTGGGGGAAGGATCACC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GAGATATTTCTTCCTTGCAACAAGCACAGGGAGAATGCCGGAGCCCGGGGTGGGGGAAGGATCACC 780

Qy   781 DATTCACACCAGTGGAACACCTTGCTTAAGTTTGCATPAAACGGCGCAATTTGATTTGCTACAA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 DATTCACACCAGTGGAACACCTTGCTTAAGTTTGCATPAAACGGCGCAATTTTATTTGCTACAA 840

Qy   841 CGCACGCCAGAGGTTGCCCCAGCCGCCGCCACCACCCCGGTTATTAGATTTGATCAAGACAGCG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CGCACGCCAGAGGTTGCCCCAGCCGCCGCCACCACCCCGGTTATTAGATTTGATCANGACAGCG 900

Qy   901 TTGAGGCCCCATCCACCGCAAAAACAGGCGCTATAGGTTGTGNGCATTAGTGTGAGGCTGGCTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTGAGGCCCCATCCACCGCAAAAACAGGCGCTATAGGTTGTGNGCATTAGTGTGAGGCTGGCTG 960

Qy   961 TTTATCGCCGGACACGATACTAATCTGGCAAATCTGGGCGGCGCACTGGAGCTCAGTACTG 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TTTATTGCCGGACACGATACTAATCTGGCAAATCTGGGCGGCGCACTGGAGCTCAACTGG 1020

Qy   1021 ACGCTTCCCGGTCAGCCGGATPAACACGGCCCGCCCAGTGTGGTGAACTGGTGTTTGAACGCTGG 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGCTTCCCGGTCAGCCGGATPAACACGGCCCGCCCAGTGTGGTGAACTGGTGTTTGAACGCTGG 1080

Qy   1081 CGTCCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGTGGTCTTCCAGACTTTTACAG 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CGTCCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGTGGTCTTCCAGACTTTTACAG 1140

Qy   1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACCGCGCCGCCGGAGAGGTGAAACTGACC 1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACCGCGCCGCCGGAGAGGTGAAACTGACC 1200

Qy   1201 CTGGGCAGGATGTGAAGAGCGAAAATGCGGCAGGGCAGAAATGTGGTCAGTGTTTTTACGGCAA 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGGCAGGATGTGAAGAGCGAAAATGCGGCAGGGCAGAAATGTGTTCGTTCGTTGGCGCAGGTTTTTACGGCAA 1260

Qy   1261 ATCGGTGAATGAAGCACGCCAATACCGCGGCGTGCAGGTTTGAGATCTCATC 1306
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAGCACGCCAATACCGCGGCGTGCAGGTTTGAGATCTCATC 1306

RESULT 9
U00096_10
WPCOMMENT
Sequence split into 47 fragments   LOCUS U00096 Accession U00096

Fragment Name   Begin   End
U00096_00   1   110000
U00096_01   100001   210000
U00096_02   200001   310000
U00096_03   300001   410000
U00096_04   400001   510000
U00096_05   500001   610000
U00096_06   600001   710000
U00096_07   700001   810000
U00096_08   800001   910000
U00096_09   900001   1010000
U00096_10   1000001   1110000
U00096_11   1100001   1210000
U00096_12   1200001   1310000
U00096_13   1300001   1410000
U00096_14   1400001   1510000
U00096_15   1500001   1610000
U00096_16   1600001   1710000
U00096_17   1700001   1810000
U00096_18   1800001   1910000
U00096_19   1900001   2010000
U00096_20   2000001   2110000
U00096_21   2100001   2210000
U00096_22   2200001   2310000
U00096_23   2300001   2410000
U00096_24   2400001   2510000
U00096_25   2500001   2610000
U00096_26   2600001   2710000
U00096_27   2700001   2810000
U00096_28   2800001   2910000
U00096_29   2900001   3010000
U00096_30   3000001   3110000
U00096_31   3100001   3210000
U00096_32   3200001   3310000
U00096_33   3300001   3410000
U00096_34   3400001   3510000
U00096_35   3500001   3610000
U00096_36   3600001   3710000
U00096_37   3700001   3810000
U00096_38   3800001   3910000
U00096_39   3900001   4010000
U00096_40   4000001   4110000
U00096_41   4100001   4210000
U00096_42   4200001   4310000
U00096_43   4300001   4410000
U00096_44   4400001   4510000
U00096_45   4500001   4610000
U00096_46   4600001   4639675

Continuation (11 of 47) of U00096 from base 1000001 (U00096 Escherichia coli K-12 MG1655

Query Match   96.7%;   Score 1264.8;   DB 15;   Length 110000;
Best Local Similarity   98.3%;   Pred. No. 0;
Matches 1278;   Conservative 0;   Mismatches 22;   Indels 0;   Gaps 0;

Qy   1 ATGAAAAGCGATCTTAATCCCACCATTTTTATTCTCTTCGTTAACCCCGCAAATCTGCA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 39840 ATGAAAAGCGATCTTAATCCCACCATTTTTATTCTCTTCGTTAACCCCGCAAATCTGCA 39899

Qy   61 TTCGCTCAGAGTGAGCGCCGGAGCTGAAGCTGAAGCTGGAAAGTGGGGAATTCAGTGGT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 39900 TTCGCTCAGAGTGAGCGCCGGAGCTGAAGCTGAAGCTGGAAAGTGGGGAATTCAGTGGT 39959

Qy   121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGCAGGATGCATGGCCA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 39960 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGCAGGATGCATGGCCA 40019

Qy   181 ACCTGGCGCCGGTAAAACTGGGGTGAGCTGACGCACCGGCACGGCGGCGGCGGCGGCGCTC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40020 ACCTGGCGCCGGTAAAACTGGGGTTGAGCTGACGCACCGGCACGGCGGCGGCGGCGGCGCTC 40079

Qy   241 GGACATTACCTGGCGTCAGCGTCTGGGTAGCCCGGACGACGGATTGCTGCCTAAAATGTGGCTGCCCG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40080 GGACATTACCTGGCGTCAGCGTCTGGGTAGCCCGGACGACGGATTGCTGCCTAAAATGTGGCTGCCCG 40139

Qy   301 CAGTCTGGTCAGGTCGGCGGATTATTGCTGATGTCGACGCGTACCCGTAAAACAGGCGAA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40140 CAGTCTGGTCAGGTCGGCGGATTATTGCTGATGTCGACGCGTACCCGTAAAACAGGCGAA 40199

Qy   361 GCCTTCGCCGCCGCGGGTGGCACTGGACACTGGTGCAAATAACCCAGCCAGATACG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40200 GCCTTCGCCGCCGCGGGTGGCACTGGACACTGGTGCAAATAACCCAGCCAGATACG 40259

Qy   421 TCCAGTCCCGGATCCGGTTATTTATTTAATCCTCTAAAAAACTGGCGTTTTGCCAACTGGATAACGCG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40260 TCCAGTCCCGGATCCGGTTATTTATTTAATCCTCTAAAAAACTGGCGTTTTGCCAACTGGATAACGCG 40319

Qy   481 AACGTGACTGACGCGATCCTCGAGGGCAGGAGGGTCAATTGCTGACTTTACCGGCAT 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40320 AACGTGACTGACGCGATCCTCGAGGGCAGGAGGGTCAATTGCTGACTTTACCGGCAT 40379

Qy   541 TATCAAACGGGTTTCCGACATGTTGACTGGTGCTTAATTTCCGAATCAAAACTGTGC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40380 CGGCAAACGGCGTTTCCGACTGGGTGGCACTGGACACTGGACAAATCAAAACTTGTGC 40439

Qy   601 CTTAAACGTGAGAAAACGGACGTGTTCATTAACGCACGGCATTACCATCCATCGGAACTC 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40440 CTTAAACGTGAGAAAACAGGACGACAAAGCTGTTCATTAACGCACGGCATTACCATCC 40499

Qy   661 AAGGTGAGCGCGGACTGTGTCTCATTAACCGGTGCGGTTAAGCCTCGGCATCAATGCTGACG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   720

RESULT 10
AP009048_10
WPCOMMENT
Sequence split into 47 fragments   LOCUS AP009048 Accession AP009048

Fragment Name	Begin	End
AP009048_00	1	110000
AP009048_01	100001	210000
AP009048_02	200001	310000
AP009048_03	300001	410000
AP009048_04	400001	510000
AP009048_05	500001	610000
AP009048_06	600001	710000
AP009048_07	700001	810000
AP009048_08	800001	910000
AP009048_09	900001	1010000
AP009048_10	1000001	1110000
AP009048_11	1100001	1210000
AP009048_12	1200001	1310000
AP009048_13	1300001	1410000
AP009048_14	1400001	1510000
AP009048_15	1500001	1610000
AP009048_16	1600001	1710000
AP009048_17	1700001	1810000
AP009048_18	1800001	1910000
AP009048_19	1900001	2010000
AP009048_20	2000001	2110000
AP009048_21	2100001	2210000
AP009048_22	2200001	2310000
AP009048_23	2300001	2410000
AP009048_24	2400001	2510000
AP009048_25	2500001	2610000
AP009048_26	2600001	2710000
AP009048_27	2700001	2810000
AP009048_28	2800001	2910000
AP009048_29	2900001	3010000
AP009048_30	3000001	3110000
AP009048_31	3100001	3210000
AP009048_32	3200001	3310000
AP009048_33	3300001	3410000
AP009048_34	3400001	3510000
AP009048_35	3500001	3610000
AP009048_36	3600001	3710000
AP009048_37	3700001	3810000
AP009048_38	3800001	3910000
AP009048_39	3900001	4010000
AP009048_40	4000001	4110000
AP009048_41	4100001	4210000
AP009048_42	4200001	4310000
AP009048_43	4300001	4410000
AP009048_44	4400001	4510000
AP009048_45	4500001	4610000
AP009048_46	4600001	4643332

Continuation (11 of 47) of AP009048 from base 1000001 (AP009048 Escherichia coli W3110 D

Query Match   96.7%;   Score 1264.8;   DB 15;   Length 110000;
Best Local Similarity   98.3%;   Pred. No. 0;
Matches 1278;   Conservative   0;   Mismatches   22;   Indels   0;   Gaps   0;

[DNA sequence alignment blocks follow, showing Qy (query) and Db (database) rows with base positions, in two columns across the page.]

```
RESULT 11
LOCUS       AR636189                1901 bp    DNA     linear   PAT 20-APR-2005
DEFINITION  Sequence 6 from patent US 6855365.
ACCESSION   AR636189
VERSION     AR636189.1  GI:62768872
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1901)
  AUTHORS   Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B.,
            O'Donoghue,E. and Mathur,E.J.
  TITLE     Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: US 6855365-A 6 15-FEB-2005;
            Diversa Corporation; San Diego, CA
FEATURES             Location/Qualifiers
     source          1..1901
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN
```

Query Match   96.6%;   Score 1263.8;   DB 2;   Length 1901;
Best Local Similarity   98.2%;   Pred. No. 0;   Matches 1277;   Conservative   0;   Mismatches 23;   Indels   0;   Gaps   0;

```
Qy        1 ATGAAAGCGGATCTTAATCCATTTTTATCTCTTCGATTCCGTTAACCCGCCAATCTGCA 60
            ||||||||| ||||||||||||||||||| ||||||||||||||||||||||||||||||
Db      188 ATGAAAGCGGATCTTAATCCATTTTTATCTCTTCGATTCCGTTAACCCGCCAATCTGCA 247
```

```
RESULT 12
LOCUS       AR636190                1901 bp    DNA     linear   PAT 20-APR-2005
DEFINITION  Sequence 7 from patent US 6855365.
ACCESSION   AR636190
VERSION     AR636190.1  GI:62768873
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1901)
  AUTHORS   Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B.,
            O'Donoghue,E. and Mathur,E.J.
  TITLE     Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: US 6855365-A 7 15-FEB-2005;
            Diversa Corporation; San Diego, CA
FEATURES             Location/Qualifiers
     source          1..1901
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match             96.6%;  Score 1263.8;  DB 2;  Length 1901;
Best Local Similarity   98.2%;
Matches  1277;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy        1 ATGAAAAGCGATCTTAATCCCATTTTATCTCTTCGATTCCGTTAACCCGGAATCTGCA   60
            |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db      188 ATGAAAACGGATCTTAATCCCATTTTATCTCTTCGATTCCGTTAACCCGGAATCTGCA  247

Qy       61 TTCGCTCAGAGTGAGCCCGGAGCTGAAGCTGAAAGTGTGGTGTCAGTCGTGTCATGGT  120
            ||||||||||||||||| |||||||||||||||||||||||||| ||||||||||||||
Db      248 TTCGCTCAGAGTGAGCCGGGAGCTGAAGCTGAAAGTGTGGTGTTCAGTCGTGTCATGGT  307

Qy      121 GTGGCGTCTCCAACCAAGGCCACGCAACTGATCGACGGCAGCCGCCATGGCCA  180
            ||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db      308 GTGGCGTCCTCAACCAAGGCCACCGCAACTGATCGACGGCAGCCGCCATGGCCA  367

Qy      181 ACCTGCGCCGGTTAAAACTGGGGTTGGCTGTGAGCTGATCACCCGAGACGGCCATTACCTC  240
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      368 ACCTGCGCCGGTTAAAACTGGGGTTGGCTGTGAGCTGATCACCCGAGACGGCCATATCTC  427

Qy      241 GGACATTACTGGCGTCAGCGTCGGCGTAGCGGATTGCTCGCCTAAAAATGTGGCTGCGCG  300
            |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db      428 GGACATTACCAACGCGAGCGGTCGGCGTTCTGGCGGATTGCTCGGCGAAAAAGGCTGCGG  487

Qy      301 CAGTCTCGGTCAGGTCGGATTATTGCTGTCGACGACGGATCCGGTTTTTACGGCA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      488 CAGTCTCGGTCAGGTCCGGATTATTGCTGTCGACGGAGGCGTACCCGTTAAAACAGCGGAA  547

Qy      361 GCCTTCGCCGCCCGCGCGGGGCTGGCAGCCTGGCACTGACTGCAATAACGGCAGATACG  420
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      548 GCCTTCGCCGCCCGCGCGGGGCTGGCAGCCTGGCACTGACTGCAATAACCGGCAGATACG  607

Qy      421 TCCAGTCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTGCCAACTGGATAACGGCG  480
            |||||||||||| |||||||||||||||| ||||||||||||||||||||||||||||
Db      608 TCCAGTCCCGATCCGTTATTTAAATCCTCTAAAAAACTGGCGTTGCCAACTGGATAACGCG  667

Qy      481 AACGTGACTGACGCGATCCTCGAGAGGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      668 AACGTGACTGACGCGATCCTCAGCAGGAGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT  727

Qy      601 CTTAAACGTGAGAAAAACAGGACGAAAAGCTGTTCATTAACCGGCAGCATTACCATCGGAACTC  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      788 CTTAAACGTGAGAAAAACAGGACGAAAAGCTGTTCATTAACCGGCAGCATTACCATCGGAACTC  847

Qy      661 AAGGTGAGCGGCCGGACCTGTGTCTCATTAACCGGTGCGCGGTAAGCCTTCGCATCAAATGCTGACG  720
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      848 AAGGTGAGCGGCCGGACCTGTGTCTCATTAACCGGTGCGCGGTAAGCCTTCGCATCAAATGCTGACG  907

Qy     1141 CAGATGGCTGATAAAAACCCGCTGTCATTAAATACGCGCGTGAAACTGACC  1200
            |||||| ||||| ||||||||| ||| ||||||||||||||||||||||||
Db     1328 CAGATGCCTGATGATAAAAACCGCGCTGTCATTAAATACCGCGGTGAAACTGACC  1387

Qy     1201 CTGGGCACGGATGTGGAAGGACGGAAATGCCGGCAGGGGCATGTGTCCGTTTTTACGGCAA  1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1388 CTGGGCACGGATGTGGAAGGACGGAAATGCCGGCAGGGGCATGTGTCCGTTTTTACGGCAA  1447

Qy     1261 ATCGTGAATGAAGCACGGCATACCGGGCCGTGCCAGTTTGTGATGAT  1300
            |||||||||||||||||||||||||||||||||||||||||||
Db     1448 ATCGTGAATGAAGCACGGCATACCGGGCCGTGCCAGTTTGTGATGAT  1487
```

```
RESULT 13
LOCUS       AX356571                1901 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 6 from Patent WO0190333.
ACCESSION   AX356571
VERSION     AX356571.1  GI:18621058
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B. and
            O'Donoghue,E.
  TITLE     Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: WO 0190333-A 6 29-NOV-2001;
            DIVERSA CORPORATION (US)
FEATURES             Location/Qualifiers

Db      668 AACGTGACTGACGCGATCCTCAGCAGGAGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT  727

Qy      541 TATCAAAACGCGTTTGGCGAACTGGAACGGGTTGCTTAATTTTCCGGCAATCAAACTTGTGC  600
            ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db      908 GAGATATTTCTCTTGCGACAACAGCAGGAAAATGCCCGGAGCCGGGGGTGGGGAAGGATCACC  967

Qy      721 GAGATATTTCTCTTGCGACAACAGCAGGAAAATGCCCGGAGCCGGGGGTGGGGAAGGATCACC  780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      908 GAGATATTTCTCTTGCGACAACAGCAGGAAAATGCCCGGAGCCGGGGGTGGGGAAGGATCACC  967

Qy      781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCAATTTGATTTGCTACAA  840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      968 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCAATTTTATTTTCGCTACAA  1027

Qy      841 CGCACGCACCAGAGGTTGCCCGGCCAGCCGCCCGCCCCCCGTTATTAGATTTGATCAAGACAGCG  900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1028 CGCACGCACCAGAGGTTGCCCGGCCAGCCGCCCGCCCCCCGTTATTAGATTTGATCAAGACAGCG  1087

Qy      901 TTGACGCCCCATCCACCCCAAAAAACAGGCGTATGGTGGTGACATTACCCACTTCAGTGCTG  960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1088 TTGACGCCCCATCCACCCCAAAAAACAGGCGTATGGTGGTGACATTACCCACTTCAGTGCTG  1147

Qy      961 TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGGGCGGCCACTGGAGCTCAACTGG  1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1148 TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGGGCGGCCACTGGAGCTCAACTGG  1207

Qy     1021 ACGCTTCCCGGTGCTCAGCGCCGGATAACACGCGCCAGGTGGTGAACTGGCTGGTGAAGCGCTCG  1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1208 ACGCTTCCCGGTGCTCAGCGCCGGATAACACGCGCCAGGTGGTGAACTGGCTGGTGAAGCGCTCG  1267

Qy     1081 CGTCGGCGTAAGCGGTAGCGAATAACAGCCAGTGGATTCAGGTTTCGGTGTTTCGGGACTTACG  1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1268 CGTCGGCGTAAGCGGTAGCGAATAACAGCCAGTGGATTCAGGTTTCGGTGTTTCGGGACTTACAG  1327

Qy     1141 CAGATGCCTGATGATAAAAACCGCGCTGTCATTAAATACGCGCGGTGAAACTGACC  1200
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1328 CAGATGCCTGATGATAAAAACCGCGCTGTCATTAAATACGCGCGGTGAAACTGACC  1387

Qy     1201 CTGGGCACGGATGTGGAAGGACGGAAATGCCGGCAGGGGCATGTGTCCGTTTTTACGGCAA  1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1388 CTGGGCACGGATGTGGAAGGACGGAAATGCCGGCAGGGGCATGTGTCCGTTTTTACGGCAA  1447

Qy     1261 ATCGTGAATGAAGCACGGCATACCGGGCCGTGCCAGTTTGTGATGAT  1300
            |||||||||||||||||||||||||||||||||||||||||||||||
Db     1448 ATCGTGAATGAAGCACGGCATACCGGGCCGTGCCAGTTTGTGATGAT  1487
```

```
source          1..1901
                /organism="Escherichia coli"
                /mol_type="unassigned DNA"
                /db_xref="taxon:562"

ORIGIN

Query Match   96.6%;  Score 1263.8;  DB 2;  Length 1901;
Best Local Similarity   98.2%;  Pred. No. 0;
Matches 1277;  Conservative   0;  Mismatches  23;  Indels   0;  Gaps   0;

Qy      1 ATGAAAGCGATCTTAATCCATTTTTATCTCTTCTGATTCGGTTAACCCCGCAAATCTGCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188 ATGAAAGCGATCTTAATCCATTTTTATCTCTTCTGATTCGGTTAACCCCGCAAATCTGCA 247

Qy     61 TTCGCTCAGAGTGAGCCGGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248 TTCGCTCAGAGTGAGCCGGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 307

Qy    121 GTGGCTGCTCCAACCAAGGCCACGCAACTGATGCAGGATATTCACCCCGAGCGACGCCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308 GTGGCTGCTCCAACCAAGGCCACGCAACTGATGCAGGATATTCACCCCGAGCGACGCCA 367

Qy    181 ACCTGGCCGGGTAAAAACTGGGTTGGGCTGACACCGCCGGGGTTGAGCTAATCGGCTTATCTC 240
          |||||||||||||  || ||||||||||||||||||||||||||||||||||||||||||
Db    368 ACCTGGCCGGGTAAAAACTGGGTTGGGCTGACACCGCCGGGNGGTTGAGCTAATCGGCTTATCTC 427

Qy    241 GGACATTACTGGCGTCAGCGGTCTGGTAGCCGACGGATTGCTGCCTAAATGTGGCTGCCCG 300
          ||||||||||||||||||||||||||||||||||||||||||  ||| ||||||||||||
Db    428 GGACATTACCTGGCGTCAGCGGTCTGGTAGCCGACGGATTGCTGCCGAAAAGGCTGCCCG 487

Qy    301 CAGTCTTGGTCAGGTCGCGATTATTGCTGATTGTCGACGGCGTACCCGTAAAACAGGCGAA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488 CAGTCTTGGTCAGGTCGCGATTATTGCTGATTGTCGACGGCGTACCCGTAAAACAGGCGAA 547

Qy    361 GCCTTCGCGCCGGGCCGGGCGGGCCCACCTGACTGTGTCAATAACCGTACATACCCAGACGACTACG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548 GCCTTCGCGCCGGGCCGGGCGGGCCCACCTGACTGTGTCAATAACCGTACATACCCAGACGACTACG 607

Qy    421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGGGTTGCCAACTGGATAACGCG 480
          |||||||||||||||||||||||||||||||||||| |||  |||||||||||||||||
Db    608 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGGGTTTGCCAACTGGATAACGCG 667

Qy    481 AACGTGACTGACGCGATCCCTCGAGAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    668 AACGTGACTGACGCGATCCCTCGAGAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 727

Qy    541 TATCAAAACGGCGTTTGGCGAACTGGAACCGGGTGCTTAATTTTCCGCAAATCAAAACTTGTGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    728 CGGCAAAACGGCGTTTGGCGAACTGGAACCGGGTGCTTAATTTTCCGCAAATCAAAACTTGTGC 787

Qy    601 CTTAAACGTGAGAAACAGGACGAAAGCGTCGTTCATTAAGCAGCCATTACCATCGGAACTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    788 CTTAAACGTGAGAAACAGGACGAAAGCTGTGTTCATTAAGCGAGCCATTACCATCGGAACTC 847

Qy    661 AAGGTGAGCGCCGACTGTGTCTCATTAACCGGTGGTAACCTCGCATCAAATGTCTGACG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    848 AAGGTGAGCGCCGACAAATGTCTCATTAACCGGTGGTAAGCCTCGCATCAAATGCTGACG 907

Qy    721 GAGATATTTCTCCCTGCAACAAGCACAGGGAATGCCCGGAGCCGGGAGGAGGATCACC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908 GAGATATTTCTCCCTGCAACAAGCACAGGGAATGCCCGGAGCCGGGAGGAGGATCACC 967

Qy    781 GATTCACACAGTGGAACACAGGTTGCTAAGTTTGCATAACGGCCAATTGATTGCTACAA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968 GATTCACACAGTGGAACACACTTGCTAAGTTTGCATAACGGCCAATTGATTGCTACAA 1027

Qy    841 CGCAGCCCAGAGGTTGCCCGCAGCCCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028 CGCACGCCAGAGGTTGCCCGCAGCCCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 1087

Qy    901 TTGACGCCCCCATCCACCGCAAAAAACAGGCGGCGTATGGTGGTGTGGACATTACCCACTTCAGTGCTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 TTGACGCCCCCATCCACCGCAAAAAACAGGCGGCGTATGGTGGTGTGGACATTACCCACTTCAGTGCTG 1147
```

RESULT 14
```
LOCUS       AX356572                 1901 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 7 from Patent WO0190333.
ACCESSION   AX356572
VERSION     AX356572.1  GI:18621059
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B. and
            O'Donoghue,B.
  TITLE     Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: WO 0190333-A 7 29-NOV-2001;
            DIVERSA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..1901
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN
```

Query Match   96.6%; Score 1263.8; DB 2; Length 1901;
Best Local Similarity   98.2%; Pred. No. 0;
Matches 1277; Conservative   0; Mismatches 23; Indels   0; Gaps   0;

```
Qy      1 ATGAAAGCGATCTTAATCCATTTTTATCTCTTCTGATTCGGTTAACCCCGCAAATCTGCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188 ATGAAAGCGATCTTAATCCATTTTTATCTCTTCTGATTCGGTTAACCCCGCAAATCTGCA 247

Qy     61 TTCGCTCAGAGTGAGCCGGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248 TTCGCTCAGAGTGAGCCGGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 307

Qy    121 GTGGCTGCTCCAACCAAGGCCACGCAACTGATGCAGGATATTCACCCCGAGCGACGCCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308 GTGGCTGCTCCAACCAAGGCCACGCAACTGATGCAGGATATTCACCCCGAGCGACGCCA 367

Qy    181 ACCTGGCCGGGTAAAAACTGGGTTGGGCTGACACCGCCGGGGTTGAGCTAATCGGCTTATCTC 240
          |||||||||||||  || ||||||||||||||||||||||||||||||||||||||||||
Db    368 ACCTGGCCGGGTAAAAACTGGGTTGGGCTGACACCGCCGGGNGGTTGAGCTAATCGGCTTATCTC 427

Qy    241 GGACATTACTGGCGTCAGCGGTCTGGTAGCCGACGGATTGCTGCCTAAATGTGGCTGCCCG 300
          ||||||||||||||||||||||||||||||||||||||||||  ||| ||||||||||||
Db    428 GGACATTACCAAGGCGTCAGCGGTCTGGTAGCCGACGGATTGCTGCCGAAAAAGGCTGCCCG 487
```

```
RESULT 15
ECOAPPAA
LOCUS       ECOAPPAA              1901 bp    DNA     linear   BCT 04-APR-2002
```

```
Qy    301 CAGTCTTCGGTCAGGTCGGCGGATTATTGCTGAGTGGTGGCGGTACCCGTAAAACAGGCGGAA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488 CAGTCTTCGGTCAGGTCGGCGGATTATTGCTGAGTGGTGGCGGTACCCGTAAAACAGGCGGAA 547

Qy    361 GCCTTCGGCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548 GCCTTCGGCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 607

Qy    421 TCCAGTCCGGACCGGATTTATTTAATCCCCTAAAAACTGGCGTTTTGGCGACTGATTAACGCG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608 TCCAGTCCGGATCCGTTATTTAATCCCTCTAAAAACTGGCGTTTTGGCGACTGATTAACGCG 667

Qy    481 AACGTGACTGACGGCGGATCCTCGAGAGGGGGTCAATTGGCTACATTACCCAGGGGGCAT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    668 AACGTGACTGACGGCGATCCTCAGCAGCGGGGTCAATTGCTGACTTTACCGGGGCAT 727

Qy    541 TATCAAAAGGGCGTTTCGGGAACTGGAACGGGTGGCCTTAATTTCCGGCAATCAAAACTTGTGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    728 CGGCAAACGGGCGTTTCGGGAACTGGAACGGGTGGCCTTAATTTCCGGCAATCAAAACTTGTGC 787

Qy    601 CTTAAACGTGACGAAACAGGACGAAAAGCTGTTCATTATTAACGCGCAGCATTACCATCGGAACTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    788 CTTAAACGTGAGAAAACAGGACGAAAAGCTGTTCATTATTAACGCGCAGCATTACCATCGGAACTC 847

Qy    661 AAGGTGAGCGGCCGACTGTGTCTCATTAACCTTAATTTCCGGCAATCAATGCTGACG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    848 AAGGTGAGCGGCCGACAAATGTCTCATTAACCGGTGCGGTAAGCCTCGGCATCAATGCTGACG 907

Qy    721 GAGATATTTCTCCTGCAACAAGCACAGGGAATCCGGAGGCCGGGTGGGGAAGGATCACC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCCGGGTGGGGAAGGGATCACC 967

Qy    781 GATTCACACAGTGGAACACCTTGCTAAGTTTGCATAACGGCGCCAATTGATTGCTACAA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968 GATTCACACAGTGGAACACCTTGCTAAGTTTGCATAACGGCGCCAATTTATTGCTACAA 1027

Qy    841 CGCACGCCAGAGGTTGCCCGGCCACCCCGTTAGTTATTGAGATTTGAACGAGACAGCG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028 CGCACCCCGAGAGGTTGCCCGGCCACCCCGTTATTTAGATTTGATCAAGAGACAGCG 1087

Qy    901 TTGACGCCCATCCACCGCAAAAAACAGGCGTATTACCCCACTTCAGTGCTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 TTGACGCCCCATCCACCGCAAAAAACAGGCGTATGCGGTTGACATTACCCCACTTCAGTGCTG 1147

Qy    961 TTTATCGCCGGACACACGATACTAATTCTCGGCAAATCTCGGCGCCAGAGCTCAACTGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148 TTTATCGCCGGACACGATACTAATCTCGCAAATCTCGGCGGCGCGGGAGCTCAACTGG 1207

Qy   1021 ACGCTTCCCGGTCAGCGGGATAACACCCCGGCCAGGGCGTGGTGAACTGGTGTTTGAACGGCTGG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208 ACGCTTCCCGGTCAGCGCGGATAACACCCCGGCCAGGGCGTGGTGAACTGGTGTTTGAACGGCTGG 1267

Qy   1081 CGTCGGCTAAGGGATAACAGCCAGTGGATTCAGGGTTTGCTCGTCTTTCCAGACTTTTACAG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268 CGTCGGCTAAGGGATAACAGCCAGTGGATTCAGGGTTTGCTCGGTTCTTTCCAGACTTTTACAG 1327

Qy   1141 CAGATGCCGTGATAAAAACGCCCGTCATTAAATAACCGCGCGGAGAGGTGAAAACTGACC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328 CAGATGCCGTGATAAAAACGCCCGTCATTAAATAACCGCGCCGGAGAGGTGAAAACTGACC 1387

Qy   1201 CTGGCAGGAGTGTGAACGGCGAAATGCCCAGGGCCATGTGTTCGTTGGGCAGGTTTACGCAA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388 CTGGCAGGATCTGAAGAGCGAAATGCCCAGGGCCATGTGTTCGTTGGGCAGGTTTACGCAA 1447

Qy   1261 ATCGTGAATGAAGCACGCCATACCGGGCGTGCAGTTTGAGAT 1300
          ||||||||||||||||||||||||||||||| ||
Db   1448 ATCGTGAATGAAGCACGCCATACCGGGCGTGCAGTTTGTTAAT 1487
```

```
FEATURES             Location/Qualifiers
     source          1..1901
                     /organism="Escherichia coli"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:562"
     gene            3..1522
                     /gene="appA"
     RBS             3..8
                     /gene="appA"
     CDS             16..108
                     /gene="appA"
                     /note="ORFX"
                     /codon_start=1
                     /transl_table=11
                     /product="unknown"
                     /db_xref="GI:145284"
                     /protein_id="AAA72085.1"
                     /translation="MWYLLWFVGILLMCSLSSTLVLVWLDPRLKS"
     stem_loop       118..149
                     /gene="appA"
     -35_signal      133..138
                     /gene="appA"
     -10_signal      156..161
                     /gene="appA"
     RBS             172..177
                     /gene="appA"
     CDS             188..1486
                     /gene="appA"
                     /standard_name="acid phosphatase"
                     /codon_start=1
                     /transl_table=11
                     /product="periplasmic phosphoanhydride phosphohydrolase"
                     /protein_id="AAA72086.1"
                     /db_xref="GI:145285"
                     /translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP
                     TKATQLMQDVTPDAWPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQS
                     GQVAIIADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA
                     NVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPS
                     ELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQF
                     YLLQRTPEVARSRATPLLDLIKTALTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGG
                     ALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNT
                     PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL"
     sig_peptide     188..253
                     /gene="appA"
```

```
mat_peptide    254..1483
               /gene="appA"
               /product="periplasmic phosphoanhydride phosphohydrolase"
variation      299..301
               /gene="appA"
               /standard_name="H303A mutant"
               /note="created by site directed mutagenesis"
               /citation=[2]
variation      299..301
               /gene="appA"
               /standard_name="R16A mutant"
               /note="created by site directed mutagenesis"
               /citation=[2]
variation      299..301
               /gene="appA"
               /standard_name="R20A mutant"
               /note="created by site directed mutagenesis"
               /citation=[2]
variation      299..301
               /gene="appA"
               /standard_name="R92A mutant"
               /note="created by site directed mutagenesis"
               /citation=[2]
variation      302..304
               /gene="appA"
               /standard_name="H17N mutant"
               /note="created by site directed mutagenesis"
               /citation=[2]
variation      1163..1165
               /gene="appA"
               /standard_name="D304A mutant"
               /note="created by site directed mutagenesis"
               /citation=[3]
               /phenotype="Asp 304 changed to Ala"
               /replace="gct"
variation      1163..1165
               /gene="appA"
               /standard_name="D304E mutant"
               /note="created by site directed mutagenesis"
               /citation=[3]
               /phenotype="Asp 304 changed to Glu"
               /replace="gag"
terminator     1497..1522
               /gene="appA"
               /note="rho-independent terminator"

ORIGIN

Query Match      96.6%;  Score 1263.8;  DB 15;  Length 1901;
Best Local Similarity  98.2%;  Pred. No. 0;
Matches 1277;  Conservative  0;  Mismatches  23;  Indels  0;  Gaps  0;

Qy        1 ATGAAAGCGATCTTAATTCCATTTTATTCTCTTCGATTCCGTTAACCCCGCAATCTGCA   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      488 CAGTCTGGTCAGGTCGCGGATTATTGCTGATGTCGACGAGGGTACCCGTAAAACAGGCGAA  547

Qy      188 ATGAAAGCGATCTTAATCCCATTTTATCTCTTCGATTCCGTTAACCCCGCAATCTGCA     247
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 GCCTTCGCCGCGCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG  420

Qy      188 ATGAAAGCGATCTTAATCCCATTTTATCTCTTCGATTCCGTTAACCCCGCAATCTGCA     247
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      548 GCCTTCGCCGCGCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG  607

Qy       61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAGTGTGGTGATTGTCACCCCGCAATCTGCA  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 TCCAGTCCCGATCCGTTTATTTTAATCCTCTAAAAACTGGGTTGCCAACTGGATAACGCG  480

Qy      248 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGTGGTGATTGTCAGTCGTCATGGT        307
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      608 TCCAGTCCCGATCCGTTTATTTTAATCCTCTAAAAACTGGGTTGCCAACTGGATAACGCG  667

Qy      121 GTGGCGTGCTCCAACCAAGGCCACGCAACTGATGACGAGTGTCACCCCGAGGCATGGCCA  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 AACGTGACTGACGCGGATCCTCGAGAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT  540

Qy      308 GTGGCGTGCTCCAACCAAGGCCACGCAACTGATGACGAGTGTCACCCCGAGGCATGGCCA  367
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      668 AACGTGACTGACGCGGATCCTCGAGAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT  727

Qy      181 ACCTGGCCGGTAAAACTGGGTGAGCTGACAGCACCGCGCGGTGGTGGTGGCTAATCGCCTTATCTC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 TATCAAACGGGCGTTTGCGAACTGGAACGGGGTTGCTTAATTTCCGCGAAATCAAACTGTGC  600

Qy      368 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGCGGTGGTGAGCTAATCGCCTTATCTC  427
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      728 CGGCAAACGGGCGTTTGCGAACTGGAACGGGGTGCTTAATTTCCGCGAAATCAAACTGTGC  787

Qy      241 GGACAATTACTGGGGTCAGGCGGTCTGGTAGCCCAACGCAGCGGCGATTGCTGCCTRAAAGTGGGCTGCCCG  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 CTTAAACGTGAGAAAACAGGACGAAAAGCTGTTCATTAATCGCGGCGCATTACCATCGCGAACTC  660

Qy      428 GGACAATTACCAAACGCCAGCGRYCTCGGGTRAGCCACGCGAGGCATTGCTCGGGCACAGGGGCTGCCCG  487
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      788 CTTAAACGTGAGAAAACAGGACGAAAAGCTGTTCATTAATCGCGGCGCATTACCATCGCGAACTC  847

Qy      301 CAGTCTGGTCAGGTCGCGGATTATTGCTGATGTCGACGAGGGTACCCGTAAAACAGGCGAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      488 CAGTCTGGTCAGGTCGCGGATTATTGCTGATGTCGACGAGGGTACCCGTAAAACAGGCGAA  547
```

Db   488 CAGTCTGGTCAGGTCGCGGATTATTGCTGATGTCGACGAGGGTACCCGTAAAACAGGCGAA 547

Qy   361 GCCTTCGCCGCGCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420
Db   361 GCCTTCGCCGCGCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420

Qy   548 GCCTTGCGCCGGATAACCAGGCCAGTGGATTCAGGTTTCGCGTTGTGTTTCCAGACTTTACAG 607

Qy   421 TCCAGTCCCGATCCGTTTATTTTAATCCTCTAAAAACTGGGTTGCCAACTGGATAACGCG 480
Db   421 TCCAGTCCCGATCCGTTTATTTTAATCCTCTAAAAACTGGGTTGCCAACTGGATAACGCG 480

Qy   548 CGTCGGCTAAGGCGGATAACAGGCCAGTGGGATTCAGGTTTCGCGTTGTGTTTCCAGACTTTACAG 607

Qy   608 TCCAGTCCCGATCCGTTTATTTTAATCCTCTAAAAACTGGGTTGCCAACTGGATAACGCG 667

Qy   481 AACGTGACTGACGCGGATCCTCGAGAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 540
Db   481 AACGTGACTGACGCGGATCCTCGAGAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 540

Qy   668 AACGTGACTGACGCGGATCCTCGAGAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 727

Qy   601 CTTAAACGTGAGAAAACAGGACGAAAAGCTGTTCATTAATCGCGGCGCATTACCATCGCGAACTC 660
Db   601 CTTAAACGTGAGAAAACAGGACGAAAAGCTGTTCATTAATCGCGGCGCATTACCATCGCGAACTC 660

Qy   728 CGGCAAACGGGCGTTTGCGAACTGGAACGGGGTGCTTAATTTCCGCGAAATCAAACTGTGC 787

Qy   541 TATCAAACGGGCGTTTGCGAACTGGAACGGGGTTGCTTAATTTCCGCGAAATCAAACTGTGC 600

Qy   661 AAGGTGAGCGGCCGGACTGTGTCTTCATTAACCCGGTGCGGTGCGCTTCCGCATTACCGGAAATCGCTGACG 720
Db   661 AAGGTGAGCGGCCGGACTGTGTCTTCATTAACCCGGTGCGGTGCGCTTCCGCATTACCGGAAATCGCTGACG 720

Qy   848 AAGGTGAGCGGCCGGACAATGTCTCATTAACCCGGTGCGGTGCGCTTCCGCATTACCGGAAATCGCTGACG 907

Qy   721 GAGAGTATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCCGGGCGTGGGGAAGGGATCACC 780
Db   721 GAGAGTATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCCGGGCGTGGGGAAGGGATCACC 780

Qy   908 GAGAGTATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCCGGGCGTGGGGAAGGGATCACC 967

Qy   781 GATTCACACCAGTGGAACACCTTGCTAAAGTTTGCATAAACGGCGCAATTGATTTGCTACAA 840
Db   781 GATTCACACCAGTGGAACACCTTGCTAAAGTTTGCATAAACGGCGCAATTGATTTGCTACAA 840

Qy   968 GATTCACACCAGTGGAACACCTTGCTAAAGTTTGCATAAACGGCGCAATTGATTTGCTACAA 1027

Qy   841 CGCACCGCCAGAGGTTGCCCGCGCAAAAACAGGGTATGGTGAGACATTACCCACTTCAGTTGCTG 900
Db   841 CGCACCGCCAGAGGTTGCCCGCGCAAAAACAGGGTATGGTGAGACATTACCCACTTCAGTTGCTG 900

Qy   1088 CGCACCGCCAGAGGTTGCCCGCGCAAAAACAGGGTATGGTGAGACATTACCCACTTCAGTTGCTG 1147

Qy   961 TTTTATCGCCGGACACCGATACTAATCTGGCAAAATCTCCGGCCGCACTGGAGCTCAACTGG 1020
Db   961 TTTTATCGCCGGACACCGATACTAATCTGGCAAAATCTCCGGCCGCACTGGAGCTCAACTGG 1020

Qy   1148 TTTTATCGCCGGACACCGATAACTAATCTGGCAAAATCTCCGGCCGCACTGGAGCTCAACTGG 1207

Qy   1021 ACGCTTCCCGGTCAGCGCGGATAAACAGCGCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1080
Db   1021 ACGCTTCCCGGTCAGCGCGGATAAACAGCGCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1080

Qy   1208 ACGCTTCCCGGTCAGCGCGGATAAACAGCGCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1267

Qy   1081 CGTCGGCTAAGGCGGATAACAGCCAGTGGGATTCAGGTTTCGCGTTGTGTTTCCAGACTTTACAG 1140

Qy   1268 CGTCGGCTAAGGCGGATAACAGCCAGTGGGATTCAGGTTTCGCGTTGTGTTTCCAGACTTTACAG 1327

Qy   1141 CAGATCGTGATPAAAACGCCCCTGTCATTAAATACGGCGCAGTGGATTCAGGTGGTGAAACTGAC 1200
Db   1141 CAGATCGTGATPAAAACGCCCCTGTCATTAAATACGGCGCAGTGGATTCAGGTGGTGAAACTGAC 1200

Qy   1328 CAGATGCGTGATPAAAACGCCCGTGTCATTAAATACGGCGCAGTGGATTCAGGTGGTGAAACTGACC 1387

Qy   1201 CTGGCAGGATGTGAAGAGAGGCAGAAATGGGCAGGGCCATGTGTTCGTTGGCAGGTTTTAGGCAA 1260
Db   1201 CTGGCAGGATGTGAAGAGAGGCAGAAATGGGCAGGGCCATGTGTTCGTTGGCAGGTTTTAGGCAA 1260

Qy   1388 CTGGCAGGATGTGAAGAGAGGCAGAAATGGGCCGCAGGGCCATGTGTTCGTTGGCAGGTTTTAGGCAA 1447

Qy   1261 ATCGTGAATGAAGAGCACGCGCATGCCGGCTGCAGTTTGAGAT 1300

Qy   1448 ATCGTGAATGAAGAGCACGGCCATACCGGGCGTGCAGTTTGTAAT 1487

Search completed: June 14, 2006, 14:10:05
Job time : 7319.05 secs

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2006 Biocceleration Ltd.

GenCore version 5.1.9

Run on: June 14, 2006, 10:04:50 ; Search time 260.506 Seconds
             (without alignments)
             9394.823 Million cell updates/sec

Title: US-10-601-319-9
Sequence: 1 atgaaagcgatcttaatccc.............gcagttgagatccaatcta 1308
Perfect score: 1308

Scoring table: IDENTITY_NUC
             Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued Patents NA:*
         1: /EMC_Celerra_SIDS3/ptodata/2/ina/1_COMB.seq:*
         2: /EMC_Celerra_SIDS3/ptodata/2/ina/5_COMB.seq:*
         3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A_COMB.seq:*
         4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B_COMB.seq:*
         5: /EMC_Celerra_SIDS3/ptodata/2/ina/7_COMB.seq:*
         6: /EMC_Celerra_SIDS3/ptodata/2/ina/H_COMB.seq:*
         7: /EMC_Celerra_SIDS3/ptodata/2/ina/PCTUS_COMB.seq:*
         8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP_COMB.seq:*
         9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE_COMB.seq:*
       10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1265	96.7	1323	3	US-09-259-214-1	Sequence 1, Appli
2	1265	96.7	1323	3	US-09-318-528-1	Sequence 1, Appli
3	1265	96.7	1323	3	US-09-291-931-1	Sequence 1, Appli
4	1265	96.7	1323	3	US-09-580-515-1	Sequence 1, Appli
5	1265	96.7	1323	3	US-09-866-379-1	Sequence 1, Appli
6	1263.8	96.6	1901	3	US-09-866-379-6	Sequence 6, Appli
7	1263.8	96.6	1901	3	US-09-866-379-7	Sequence 7, Appli
8	1262.2	96.5	1901	3	US-09-866-379-5	Sequence 5, Appli
9	1262.2	96.5	1901	3	US-09-866-379-9	Sequence 9, Appli
10	1233.6	94.3	1486	3	US-09-715-477-2	Sequence 2, Appli
11	1233.6	94.3	1486	3	US-09-715-477-4	Sequence 4, Appli
12	1228	93.9	1489	4	US-10-266-041A-9	Sequence 9, Appli
13	1222.4	93.5	1486	3	US-09-715-477-4	Sequence 4, Appli
14	1153	88.1	1272	2	US-08-910-798-1	Sequence 1, Appli
15	66.2	5.1	1266	3	US-09-489-039A-341	Sequence 341, App
16	39.8	3.0	505	3	US-09-621-976-15639	Sequence 15639, A
17	36.6	2.8	2868	3	US-09-710-794-4	Sequence 4, Appli
18	35.8	2.7	1506	3	US-09-902-540-3515	Sequence 3515, Ap
19	35.8	2.7	16720	3	US-09-902-540-1168	Sequence 1168, Ap
20	34.4	2.6	4411529	3	US-09-103-840A-2	Sequence 2, Appli
21	34.4	2.6	4411529	3	US-09-103-840A-1	Sequence 1, Appli
22	34.2	2.6	4403765	3	US-09-103-840A-2	Sequence 2, Appli
23	34.2	2.6	4411529	3	US-09-103-840A-1	Sequence 1, Appli
24	34	2.6	601	3	US-09-949-016-109215	Sequence 109215,
25	34	2.6	601	3	US-09-949-016-109216	Sequence 109216,
26	34	2.6	601	3	US-09-949-016-109217	Sequence 109217,
27	34	2.6	998	2	US-07-885-089B-5	Sequence 5, Appli
28	33.8	2.6	154600	3	US-09-621-976-14757	Sequence 14757, A
29	33.6	2.6	399	3	US-09-860-200F-1	Sequence 1, Appli
30	33	2.5	1359	3	US-09-949-016-8976	Sequence 8976, Ap
31	33	2.5	2511	3	US-08-680-326-116	Sequence 116, App
32	33	2.5	69701	3	US-08-757-669A-17	Sequence 17, Appl
33	33	2.5	73308	3	US-08-770-379-17	Sequence 17, Appl
34	32.6	2.5	35100	3	US-09-230-371A-17	Sequence 17, Appl
35	32.6	2.5	35100	3	US-09-949-016-14187	Sequence 14187, A
36	32.6	2.5	35100	3	US-09-949-016-16326	Sequence 16326, A
37	32.6	2.5	3147	2	US-08-441-430-1	Sequence 1, Appli
38	32.6	2.5	4485	2	US-08-441-430-3	Sequence 3, Appl
39	32.6	2.5	4488	2	US-09-023-655-1286	Sequence 1286, Ap
40	32.6	2.5	4567	3	US-09-949-016-20	Sequence 20, Appl
41	32.6	2.5	5877	3	US-10-152-886-94	Sequence 94, Appl
42	32.4	2.5	97221	3	US-09-949-016-12755	Sequence 12755, A
43	32.2	2.5	2430	3	US-09-489-039A-3732	Sequence 3732, Ap
44	32.2	2.5	3076	3	US-09-710-794-1	Sequence 1, Appli
45	32.2	2.5	12257	3	US-09-949-016-16384	Sequence 16384, A

ALIGNMENTS

RESULT 1
US-09-259-214-1
; Sequence 1, Application US/09259214A
; Patent No. 6110719
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: DIVER1370-1
; CURRENT APPLICATION NUMBER: US/09/259,214A
; CURRENT FILING DATE: 1999-03-01
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; NAME/KEY: misc_feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-259-214-1

Query Match   96.7%;   Score 1265;   DB 3;   Length 1323;
Best Local Similarity   98.0%;
Matches 1280;   Conservative   0;   Mismatches 26;   Indels   0;   Gaps   0;

Qy   1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCGCCAATCTGCA 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCGCCAATCTGCA 60

Qy   61 TTCGCTCAGAGTGAGCCCGGAGTGGAAGTGTGGAAAGTGTGGATTGTCAGTCGTCATGGT 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCCGGAGTGGAAGTGTGGAAAGTGTGGATTGTCAGTCGTCATGGT 120

Qy   121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy   181 ACCTGGCCGGTAAAACTGGGTGAGCTGACACCCGCGGGTTGGGTGAGCTAATCGCTATCTC 240
         ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGGGTGAGCTAATCGCTATCTC 240

RESULT 2
US-09-318-528-1
; Sequence 1, Application US/09318528
; Patent No. 6183740
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/318,528
; CURRENT FILING DATE: 1999-05-25
; EARLIER APPLICATION NUMBER: 09/291,931
; EARLIER FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; NAME/KEY: misc_feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-318-528-1

Query Match   96.7%;   Score 1265;   DB 3;   Length 1323;
Best Local Similarity   98.0%;   Pred. No. 0;   Mismatches   26;   Indels   0;   Gaps   0;
Matches 1280; Conservative   0;   Mismatches   26;   Indels   0;   Gaps   0;

```
Qy    1 ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTCTTCGATTCCTTAACCCGCCAATCTGCA   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTCTTCGATTCCTTAACCCGCCAATCTGCA   60

Qy   61 TTCGCTCAGAGTGAGCCCGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCCGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT  120

Qy  121 GTGCGTGCTCCAACCAAGGCCACGCAAACTGATGTGCAGGATGATGCCCTGAGCGCCA  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTGCGTGCTCCAACCAAGGCCACGCAAACTGATGTGCAGGATGATGCCCTGAGCGCCA  180

Qy  181 ACCTGGCCGGTAAAAACTGGGTGAGCTGACACCGCGCCGTGGTAATCGCCTATCTC  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ACCTGGCCGGTAAAAACTGGGTGAGCTGACACCGCGCCGTGGTAATCGCCTATCTC  240

Qy  241 GGACATTACTGGCGTCAGCGTCGGTAGCCGACGGAGTTGCTGGCGAAAAAGGGCTGCCCG  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GGACATTACCAACGCCAGCGTCGGTAGCCGACGGAGTTGCTGGCGAAAAAGGGCTGCCCG  300

Qy  301 CAGTCTGGTCAGGTCCCGATTATTGCTGGCACCTGACCTGTGAAATAACCGTACATACCC  360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGTCTGGTCAGGTCCCGATTATTGCTGGCACCTGACCTGTGAAATAACCGTACATACCC  360

Qy  361 GCCTTCGCCGCGCGGGGCTGGGCACCTGACCTGTGAATGCTGGCGTTGCCAACTGGATACG  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GCCTTCGCCGCGCGGGGCTGGGCACCTGACCTGTGAATGCTGGCGTTGCCAACTGGATACG  420

Qy  421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAAACTGGCGTTTGCCAACTGGATAACGCG  480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAAACTGGCGTTTGCCAACTGGATAACGCG  480

Qy  481 AACGTGACTGACGCGGATCCTCAGCAGGGGCAGGAGGGGTCAATTGCTGACTTTACCGGGCAT  540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 AACGTGACTGACGCGGATCCTCAGCAGGGGCAGGAGGGGTCAATTGCTGACTTTACCGGGCAT  540

Qy  541 TATCAAACGGGCGCTTTCCGGAACTGGAACGGGTTGCTTAATTTTCCGGCAATCAAAACTTGTGC  600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 CGGCAAACGGGCGCTTTCCGGAACTGGAACGGGTTGCTTAATTTTCCGGCAATCAAAACTTGTGC  600
```

RESULT 3
US-09-291-931-1
; Sequence 1, Application US/09291931A
; Patent No. 6190897
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/291,931A
; CURRENT FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)

; NAME/KEY: misc_feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-291-931-1

Query Match 96.7%; Score 1265; DB 3; Length 1323;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```
Qy      1 ATGAAAGCGATCTTAAATCCCATATTTTATCTTCTTCTGAATTCCGTTAAACCCGCAATCTGCA    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAGCGATCTTAAATCCCATATTTTATCTTCTTCTGAATTCCGTTAAACCCGCAATCTGCA    60

Qy     61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGGAAAGTGTGGTTATTCAGTTCGTTCGTCATGGT   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGGAAAGTGTGGTTATTCAGTTCGTTCGTCATGGT   120

Qy    121 GTGGGTGCTCCAACCAAGGCCACCAGCCAACTGATGCAGGATGTCACCCCAGACGCANTGGCCA   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGGGTGCTCCAACCAAGGCCACCAGCCAACTGATGCAGGATGTCACCCCAGACGCANTGGCCA   180

Qy    181 ACCTGGCCCGGTAAAATCTGGGTGAGCTGACACCGCGCGGTGGGTGGTGAGCTAATCGCCTATCTC   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACCTGGCCCGGTAAAACTGGGTTGGCTGACACCGCGGTGGGTGGTGAGCTAATCGCCTATCTC   240

Qy    241 GGACATTACTGGCGCTCAGCGGTTCGTGGTAGCCGACGGATTGCCTGCCTAAAATGTGGCTGCCCG   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGACATTACCAACGCCAGCGGTTCGTGGTAGCCGACGGATTGCCTGCCGGAAAAAGGGCTGCCCG   300

Qy    301 CAGTCGTTGGTCAGGTCGGCGGATTATTGCTGAGTGCGACGACCCCGTAAAACAGGCGGAA   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGTCGTTGGTCAGGTCGGCGGATTATTGCTGAGTGCGACGACCCCGTAAAACAGGCGGAA   360

Qy    361 GCCTTCGCGCGGCGCTGGGCACCTGACCTGTGCAATAACCGTACTACGCAGCAGATACG   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCTTCGCGCGGCGGCGCTGGGCACCTGACTGTGCAATAACCGTACTACGCAGCAGATACG   420

Qy    421 TCCAGTCCCGATCCGTTATTTAATCCCTCTAAAAAACTGGGCTTTCCAACTGGATAAACGCG   480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCCAGTCCCGATCCGTTATTTAATCCCTCTAAAAAACTGGGCTTTCCAACTGGATAAACGCG   480

Qy    481 AACGTGACTGACCGGATCCTTCGAGCACGAGAGGGTCAATTGCTGACTTTACCGGGCAT   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AACGTGACTGACCGGATCCTTCGAGCACGAGAGGGTCAATTGCTGACTTTACCGGGCAT   540

Qy    541 TATCAAACGCGGTTTCGGGAACTGGAACGGGGTGCTTAATTTCCGCAATCAAACTTGTGC   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CGGCAAACGGCGTTTCGGGAACTGGAACGGGGTGCTTAATTTCCGCAATCAAACTTGTGC   600

Qy    601 CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCATTAACGCAGGCAGCATTACCATCGGAACTC   660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCATTAACGCAGGCAGCATTACCATCGGAACTC   660

Qy    661 AAGGTGACGCGCCCGACACTGTGTCTCATTAACCGGTGCGGTAAGCCTTCGCGATCAAATGCTGACG   720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AAGGTGACGCGCCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTTCGCGATCAAATGCTGACG   720

Qy    721 GAGATATTTCTCCTGCAACAGACANCAGGGGAAATGCCGGAGCCGGGGTGGGGAAGGATCACC   780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GAGATATTTCTCCTGCAACAGACACCTTGCTAAGTTTGCATAACGCGGAAATGCCGGAGGATCACC   780

Qy    781 GATTCACACCAGTGGGACAACCTTGCTAAGTTTGCATAACGCGGAAATTTTATTTGCTACAA   840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GATTCACACCAGTGGGACAACCTTGCTAAGTTTGCATAACGCGGAAATTTTATTTGCTACAA   840

Qy    841 CGCACGCCAGAGGTTGCCGCGCACCGGCCCACCCGTTATTGAGATTTGGTGACATTACCCACTTCAGTTGCTG   900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CGCACGCCAGAGGTTGCCGCGCACCGGCCCACCCGTTATTGAGATTTGGTGACATTACCCACTTCAGTTGCTG   900

Qy    901 TTGACGCCCCATCCACCCGCAAAAACAGGGCGTATGGGTGTGACATTACCCACTTCAGTTACTG   960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTGACGCCCCATCCACCCGCAAAAACAGGGCGTATGGGTGTGACATTACCCACTTCAGTTACTG   960
```

RESULT 4
US-09-580-515-1
; Sequence 1, Application US/09580515
; Patent No. 6720014
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/580,515
; CURRENT FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
;
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-580-515-1

Query Match   96.7%; Score 1265; DB 3; Length 1323;
Best Local Similarity   98.0%; Pred. No. 0; Mismatches 26;
Matches 1280; Conservative 0; Indels 0; Gaps 0;

[Sequence alignment blocks — nucleotide sequences with Qy (Query) and Db (Database) rows, numbered from 1 to 1306/1323]

Db   1261 ATCGTGAATGAAGCACGGCATACCGGCGGTGCAGTTTGAGATCTCATC 1306

RESULT 5
US-09-866-379-1
; Sequence 1, Application US/09866379
; Patent No. 6855365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
US-09-866-379-1

; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n is any nucleotide
; NAME/KEY: CDS
; LOCATION: (1)..(1323)
; OTHER INFORMATION:

Query Match   96.7%; Score 1265; DB 3; Length 1323;
Best Local Similarity   98.0%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy   1 ATGAAAGCGGATCTTAATCCCATTTTATCTCTTCTGATTCCGTTACCCGCAAATCTGCA 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAGCGGATCTTAATCCCATTTTATCTCTTCTGATTCCGTTACCCGCAAATCTGCA 60

Qy   61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAAGTGTGGGATGGTGGTGTCAGTTCAGTTCATGGT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCGGAGCTGAAACTGGAAAAGTGTGGGATGTGGGTGTCAGTTCGTCGTGTCATGGT 120

Qy   121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATATGCAGGAGTCACCCGAGCGGCATGGCCA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATATGCAGGAGTCACCCCGAGCGGCATGGCCA 180

Qy   181 ACCTGGCCGGGTAAAACTGGGGTTGAGCTACGACTTAAATCCCTACTC 240
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ACCTGGCCGGGTAAAACTGGGGTTGAGCTGAGCTTAATCCCCTATCTC 240

Qy   241 GGACATTACCAACGCCAGCGTCTGGGTAGCCGACGGGATTGCTCGGGAAAAAGGGCTGCCCG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGACATTACCAACGCCAGCGTCTGGGTAGCCGACGGGATTGCTCGGGAAAAAGGGCTGCCCG 300

Qy   301 CAGTCTCGGTCAGGTCGCGGATTAATTGCTGAATGTGTCGACGAGCGTACCCCTAAAAACAGGCGAA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGTCTCGGTCAGGTCGCGGATTAATTGCTGAATGTGTCGACGAGCGTACCCCTAAAAACAGGCGAA 360

Qy   361 GCCTTCGCGCCGCGGGCCTGGCCACCTGACCTGTGCAATAACCGTTGCAATAACCCCAGGCCAGTACG 420
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCCTTCGCGCCGCGGGCCTGGCCACCTGACCTGTGCAATAACCGTTGCAATAACCCCAGGCCAGTACG 420

Qy   421 TCCAGTCCCGATCCGGTTATTTAATCCTCTAAAAAACTGGGGTTTGCCAACTGGATAACGCG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCCAGTCCCGATCCGGTTATTTAATCCTCTAAAAAACTGGGGTTTGCCAACTGGATAACGCG 480

Qy   481 AACGTGACTGACGGGCATCCTTCAGGAGGGGGTCAATTGCTTGACTTTACCCGGGCAT 540
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AACGTGACTGACGGGCATCCTTCAGGAGGGGGTCAATTGCTTGACTTTACCCGGGCAT 540

Qy   541 TATCAAAACGGCGGTTTCGGAACTGGAACGGGTGCTTAATTTCCCGCAATCAAACTGTGC 600
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CGGCAAAACGGCGGTTTCGGAACTGGAACGGGTGCTTAATTTCCCGCAATCAAACTGTGC 600

Qy   601 CTTAAACGTGAGAAAACAGGACGAAAAGCTGTTCATTAACGCCAGGAGTTC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTTAAACGTGAGAAAACAGGACGAAAAGCTGTTCATTAACGCCAGGAGAACTC 660

Qy   661 AAGGTGAGCGGCCCGACTGTGTCTCATTAACCGGGTCGCGGTCACATCAAATGCTGACG 720
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AAGGTGAGCGGCCCGACAATGTCTCATTAACCGGGTCGCGGTCACATCAAATGCTGACG 720

Qy   721 GAGATATTTCTCCTGCAACAAGCACAGGAAATGCCGGAGCGGGGGTGGGGAGGATCACC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GAGATATTTCTCCTGCAACAAGCACAGGAAATGCCGGAGCGGGGGTGGGGAGGATCACC 780

Qy   781 GATTCACCACCAGTGGAACACCTTCGCTAAGTTTGCATAACGCGGCAAATTTGATTTGCTACAA 840
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GATTCACCACCAGTGGAACACCTTCGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA 840

Qy   841 CGCACGCCAGAGAGGTTGCCCGCGCCCAGGTGTGGTGAACTGTGTTTGAACGCTGG 900
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CGCACGCCAGAGAGGTTGCCCGCGCCCAGGTGTGGTGAACTGGTGTTTGAACGCTGG 900

Qy   901 TTGACGCCCCATCCACCGCAAAAACAGGGGTATGGTGTGAACATTACCCACTTCAGTTACTG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTGACGCCCCATCCACCGCAAAAACAGGGGTATGGTGTGAACATTACCCACTTCAGTTACTG 960

Qy   961 TTTATCGCGGACACGGATACTAATCTGGCAAAATCTCGGCGCGCCACTGGAGCTCAACTGG 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TTTATTGCCGGACACGGATACTAATCTGGCAAATCTCGGCGCGCCACTGGAGCTCAACTGG 1020

Qy   1021 ACGCTTCCCGGTCAGCCCGGATAACACCGCGCCAGGTGGTGAACTGGTGTTTTGAACGCTGG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGCTTCCCGGTCAGCCCGGATAACACCGCGCCAGGTGGTGAACTGGTGTTTTGAACGCTGG 1080

Qy   1081 CGTGCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTCGCTGGTTCTTCCAGACTTTACAG 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CGTGCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTCGCTGGTTCTTCCAGACTTTACAG 1140

Qy   1141 CAGATGCCTGATAAAAACCCCGCGTTCATTAAAATACGGCCCGCCCCCGAGAGGTGAAAACTGACC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAGATGCCTGATAAAAACCCCGCGTTCATTAAAATACGGCCCGCCCCCGAGAGGTGAAAACTGACC 1200

Qy   1201 CTGGCAGGATGTGAAAGAGGGAAATGCCGCAGGGCCATGTGTTCGTTGGGCAGGTTTTACGGCAA 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGCAGGATGTGAAAGAGGGAAATGCCGCAGGGCCATGTGTTCGTTGGGCAGGTTTTACGGCAA 1260

Qy   1261 ATCGTGAATGAAGCACGGCATACCGGCGGTGCAGTTTGAGATCTCATC 1306
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAGCACGGCATACCGGCGGTGCAGTTTGAGATCTCATC 1306

RESULT 6
US-09-866-379-6
; Sequence 6, Application US/09866379
; Patent No. 6855365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith

```
;   APPLICANT:  GRAY, Kevin
;   APPLICANT:  BARTON, Nelson
;   APPLICANT:  GARRETT, James
;   APPLICANT:  O'DONOGHUE, Eileen
;   TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
;   FILE REFERENCE: DIVER1370-7
;   CURRENT APPLICATION NUMBER: US/09/866,379
;   CURRENT FILING DATE:  2001-05-24
;   PRIOR APPLICATION NUMBER: US 09/580,515
;   PRIOR FILING DATE: 2000-05-25
;   PRIOR APPLICATION NUMBER: US 09/318,528
;   PRIOR FILING DATE: 1999-05-25
;   PRIOR APPLICATION NUMBER: US 09/291,931
;   PRIOR FILING DATE: 1999-04-13
;   PRIOR APPLICATION NUMBER: US 09/259,214
;   PRIOR FILING DATE: 1999-03-01
;   PRIOR APPLICATION NUMBER: US 08/910,798
;   PRIOR FILING DATE: 1997-08-13
;   NUMBER OF SEQ ID NOS: 10
;   SOFTWARE: Patentin version 3.1
;
;   SEQ ID NO 6
;   LENGTH: 1901
;   TYPE: DNA
;   ORGANISM: Escherichia coli
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1901)
;   OTHER INFORMATION: n is any nucleotide

US-09-866-379-6

  Query Match       96.6%;  Score 1263.8;  DB 3;  Length 1901;
  Best Local Similarity    98.2%;  Pred. No. 0;  Mismatches  23;  Indels  0;  Gaps  0;
  Matches 1277;  Conservative    0;

Qy         1 ATGAAAGCGATCTTAATCCCATTTTTATCTTCTTCTCGATTCCGTTAACCCCCGCAATCTGCA 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       188 ATGAAAGCGATCTTAATCCCATTTTTATCTTCTTCTCGATTCCGTTAACCCCCGCAATCTGCA 247

Qy        61 TTCGCTCAGAGTGAGCCCGGAGCTGAGAAGTGTGGTGAATTGTCAGTGTCGTCATGGT 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       248 TTCGCTCAGAGTGAGCCCGGAGCTGAGAAGTGTGGTGAATTGTCAGTGTCGTCATGGT 307

Qy       121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGACGGACGATGTCACCCGAGACGGCCA 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGACGGACGATGTCACCCGAGACGGCCA 367

Qy       181 ACCTGCCCGGTAAAACTGGGGTGAGCTAGACACCGCGCGGGGTGAGCTAAATCGGCCTATCTC 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       368 ACCTGCCCGGTAAAACTGGGGTTGGGCTGACACCGCGCGGGGTGAGCTAATCGCCTATCTC 427

Qy       241 GGACATTACTGGGCGTCAGGTCGGTAGCCGACGGATTGCTCGCCTAAAGTGGCTGGCCCG 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       428 GGACATTACCTGGGCCAGGGGTCTGGTAGCCGACGGATTGCTCGCGGAAAAACAGGCTGCCGAA 487

Qy       301 CAGTCGTGGTCAGGTCCGGATTATTGCTGATGTCGACGAGCTACCCGTAAAACACGCGAA 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       488 CAGTCGTGGTCAGGTCCGGATTATTGCTGATGTCGACGAGCTTACCCGTAAAACACGCGAA 547

Qy       361 GCCTTCGCCCGCCGCCGGGCGTGGCACCCTGACTGTGCGCAATAACCGTACATACCCGGCAGATACG 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       548 GCCTTCGCCCGCCGCCGGGCGTGGCACCCTGACTGTGCGCAATAACCGTACATACCCGGCAGATACG 607

Qy       421 TCCAGTCCCGATCCGGTTATTTAATTCTCTAAAAACTGGGTTGCCAACTGGATAACGCG 480
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       608 TCCAGTCCCGATCCGGTTATTTAATCCTCTAAAAAACTGGGTTTGCCAACTGGATAACGCG 667

Qy       481 AACGTGACTGACGCGCGATCCTCGAGAGGGCAGGAGGCAGGAGGGTCAATTGCTAACTTTACCGGGCAT 540
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       668 AACGTGACTGACGCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTAACTTTACCGGGCAT 727

Qy       541 TATCAAAACGGGCGTTTCGCGAACTGGAACGGGCTGGCTTAATTTTCCGGCAATCAAACTTGTGC 600
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       728 CGGCAAAACGGGCGTTTCGCGAACTGGAACGGGCTGCTTAATTTTCCGGCAATCAAACTTGTGC 787
```

```
Qy       601 CTTAAAACGTGAGAAAACAGGAGCGAAAAGCTGTTCATTAAGCGACAGCAGGCAATTACCCATCGGGACTC 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       788 CTTAAAACGTGAGAAAACAGGAGCGAAAAGCTGTTCATTAAGCGACAGCAGGCAATTACCCATCGGGACTC 847

Qy       661 AAGGTGAGCGCCGACTGTGTCTCATTAACCGGTGCGTGGAAGCCTTCGCCATCAATGCTACACC 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       848 AAGGTGAGCGCCGACTGTGTCTCATTAACCGGTGCGTGGAAGCCTTCGCCATCAATGCTACACC 907

Qy       721 GAGATATTTCTCCTTGCAACACAGCACAGGAGAAATGCCGGAGCCGGGGTGGGGAGGGATCACC 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       908 GAGATATTTCTCCTTGCAACACAGCACAGGAGAAATGCCGGAGCCGGGGTGGGGAGGGATCACC 967

Qy       781 GATTCACACACCAGTGGGAACACCCTTGCTCAAGTTTGCATAAACGCCGCAATTTGATTTGCTACAA 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       968 GATTCACACACCAGTGGGAACACCCTTGCTCAAGTTTGCATAAACGCCCAATTTTATTTGCTACAA 1027

Qy       841 CGCACGCCACGAGGTTGCCCGCAGCCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 900
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1028 CGCACGCCACGAGGTTGCCCGCAGCCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 1087

Qy       901 TTGACGCCCCATCCACCCGCAAAAAACAGGCGTAATGGTGTTGACATTTACCCACTTCAGTGCTG 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1088 TTGACGCCCCATCCACCCGCAAAAAACAGGCGTATGGTGTTGACATTTACCCACTTCAGTGCTG 1147

Qy       961 TTTATCCGCCGGACACGGATTACTAAATCTGGCAAAATCTCGGCGCGGGCCACTGGGAGCTCAACTGG 1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1148 TTTATCCGCCGGACACGGATTACTAAATCTGGCAAAATCTCGGCGCGGGCCACTGGGAGCTCAACTGG 1207

Qy      1021 ACGCTTCCCGGTCAGCCCGGATAACACGGCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1208 ACGCTTCCCGGTCAGCGCGGATAACACCGCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1267

Qy      1081 CGTCGGCTAAGCCGGATATAACAGCCAGTGGCATTCAGGTTTCGTCGCTGGTTCTTCCCAGACTTTACAG 1140
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1268 CGTCGGCTAAGCCGGATATAACAGCCAGTGGCATTCAGGTTTCGCTGGTTTCTTCCCAGACTTTACAG 1327

Qy      1141 CAGATGCGTGATAAAAACGGCCCGCTGTCATTAAAATACGGCGACGGTTTCGGCGAAGCTGACC 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1328 CAGATGCGTGATAAAAACGGCCCGCTGTCATTAAAATACGGCGACGGTTTCGGCGAAGCTGACC 1387

Qy      1201 CTGGCAGGAGTGAAGAGGCGGAAATGCCGGAGCCACGATGTGTTCGTTGGCAGGGTTTTTACCGCAA 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1388 CTGGCAGGAGTGAAGAGGCGGAAATGCCGGAGCCACGATGTGTTCGTTGGCAGGGTTTTTACCGCAA 1447

Qy      1261 ATCGTGAATGAAGCACCGCATACCGGCGCTGCAGTTTGAGAT 1300
             |||||||||||||||||||||||||||||||||||||||
Db      1448 ATCGTGAATGAAGCACCGCATACCGGCGCTGCAGTTTGTTAAT 1487
```

```
;     PRIOR FILING DATE: 1999-03-01
;     PRIOR APPLICATION NUMBER: US 08/910,798
;     PRIOR FILING DATE: 1997-08-13
;   NUMBER OF SEQ ID NOS: 10
;   SOFTWARE: PatentIn version 3.1
;
; SEQ ID NO 7
;   LENGTH: 1901
;   TYPE: DNA
;   ORGANISM: Escherichia coli
;   FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (1)..(1901)
;     OTHER INFORMATION: n is any nucleotide
US-09-866-379-7

Query Match        96.6%;     Score 1263.8;   DB 3;   Length 1901;
Best Local Similarity   98.2%;   Pred. No. 0;   Mismatches  23;   Indels   0;   Gaps   0;
Matches 1277;  Conservative   0;   Mismatches  23;   Indels   0;   Gaps   0;
```

```
Qy      1 ATGAAAGGATCTTAATCCCATTTTATCTTCTGATTCGGTTAACCCGGCAATCTGCA 60
Db    188 ATGAAAGGGATCTTAATCCCATTTTATCTTCTGATTCGGTTAACCCGGCAATCTGCA 247

Qy     61 TTCGCTCAGAGTGAGCCGAGCTGAAGCTGGAAAGTGTGGTGATTCAGTCGTCATGGT 120
Db    248 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTCAGTCGTCATGGT 307

Qy    121 GTGCGTCGTCTCCAACCAAGCCACGCCACGCACTGATGAGTGCAGGAATGCAGGCAT GGCCA 180
Db    308 GTGCGTCGTCTCCAACCAAGCGCACGCACTGATGCAGGAATGTCAGGCACGGCAT GGGCA 367

Qy    181 ACCTGGGCCGGGTAAAACTGGGTGAGCTGACAGCTAATCCCACTATCTC 240
Db    368 ACCTGGGCCGGGTAAAACTGGGTTGGCTGAGCTAATCGCGCCTATCTC 427

Qy    241 GGACATTACTGGCGTCAGCGTCTCGGTAGCCGACGGATTGCTTGCCTGCCTGCCG 300
Db    428 GGACATTACCAACGGCCAGCGTCTGGTAGCCGACGGATTGCTTGCGAAAAACAGGCCG 487

Qy    301 CAGTCTCGGTCAGGTCGCGATTAATTGCTGAATGTCGACGGTACCCGTAAAACAGGCGGAA 360
Db    488 CAGTCTCGGTCAGGTCGCGGATTAATTGCTGAATGTCGACGGTACCCGTAAAACAGGCGGAA 547

Qy    361 GCCTTCGCGCCCGGGCTGGGCACCTGACTGTGCAATAACCGTACATACCCAGGACGATACG 420
Db    548 GCCTTCGCGCCGCGGGCTGGGCACCTGACTGTGCAATAACCGTACATACCCAGGACGATACG 607

Qy    421 TCCAGTCCCGGATCCGGTTATTTTAATCCTCTAAAAACTGGCGGTTTGCCAACTGGATAACGCG 480
Db    608 TCCAGTCCCGGATCCGGTTATTTTAATCCCTCTAAAAACTGGCGGTTTGCCAACTGGATAACGCG 667

Qy    481 AACGTGACTGACCCGATCCTCGAGAGGGGCGAGAGGTCAATTGCTGACTTACCGGGCAT 540
Db    668 AACGTGACTGACGCGGATCCTCAGCACGGGCGAGAGGTCAATTGCTGACTTACCGGGCAT 727

Qy    541 TATCAAAACGGGCGTTTCCGGAACTGGAACGGGGTGCTTATTTTCCGCAATCAAACTTGTGC 600
Db    728 CGGCAAACGGGCGTTTCCGGAACTGGAACGGGGTGCTTATTTTCCGCAATCAAACTTGTGC 787

Qy    601 CTTAAAACGTGAGAAAACAGGACGCCGAAAGTCGTTCATTAACGCAGCGCCATTACCATCGGGACTC 660
Db    788 CTTAAAACGTGAGAAAACAGGACGCCGAAAGTCGTTCATTAACGCAGCGCCATTACCATCGGGACTC 847

Qy    661 AAGGTGAGCGCCCGACTGTGTCTCATTAACCGGTGCGGGTAAGCCTGCGACG 720
Db    848 AAGGTGAGCGCCCGACAATGTCTCATTAACCGGTGCGGGTAAGCCTGCGACG 907

Qy    721 GAGATATTTCTCCTGCCACACAGCACAGGGAATCCCGGAGCCGGGTGGGGAGGATCACC 780
Db    908 GAGATATTTCTCCTGCCACACAGCACAGGGAATCCCGGAGCCGGGTGGGGAAGGATCACC 967

Qy    781 GATTCACACCAGTGGGAACACCTTGCTAAGTTTCCATAACGCGCAATTTGATTGCTACAA 840
Db    968 GATTCACACCAGTGGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTGCTACAA 1027
```

```
Db    841 CGCACGCCAGAGGTTGCCCGCGCAGCCGCCGACCACCCGCCGTTATTAGATTGATAGATGACAAGAGACGACG 900
Db    901 CGCACGCCAGAGGTTGCCCGCGCAGCCGCCGACCACCCGCCGTTATTAGATTTGATCAAGACAGCG 960
Qy    901 TTGACGCCCATCCACCGCAAAAAACAGGGGTATGTTGTGACATTACCCACTTCAGTGCTG 960
Db   1028 CGCACGCCAGAGGTTGCCCGCGCAGCCGCCGACCACCCGCCGTTATTAGATTTGATCAAGACAGCG 1087
Qy    961 TTGACGCCCATCCACCGCAAAAAACAGGGGTATGTTGTGACATTACCCACTTCAGTGCTG 1020
Db   1088 TTGACGCCCCATCCACCGCAAAAAACAGGGGTATGTTGTGACATTACCCACTTCAGTGCTG 1147
Qy   1021 ACGGCTTCCCGGTCAGCCGGATAACACACGCCCCAGGTGGTGAACTGGTGTTGTTTGGACGCTGG 1080
Db   1148 TTTATCGGCGGACGACTAATCTGGCAAAATCTCGGCGGCCGGCGGCACTGGACTCAACTGG 1207
Qy   1081 CGTCGGGCTAAGCGTAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140
Db   1208 ACGCTTCCCGGTCAGCCGGATAACACGCCCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1267
Qy   1141 CAGATGCCGTGAATAAAACGCCGCTGTCATTAAATAACGCCGGCCCGGAGAGGTTGAAACTGACC 1200
Db   1268 CGTCGGGCTAAGCGTAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1327
Qy   1201 CTGGCAGGATGTGGAAGAGCGGAAAATGCCGCAGGGCCAATGTGGTTCCGTTGGGCAGTTTTACGGCAA 1260
Db   1328 CAGATGCCGTGAATAAAACGCCGCTGTCATTAAATAACGCCGGCCCGGAGAGGTTGAAACTGACC 1387
Qy   1261 ATCGTGAATGAATGAGCACGCATACCGGGCGTGCAGTTGAGAT 1300
Db   1388 CTGGCAGGATGTGGAAGAGCGGAAAATGCCGCAGGGCCAATGTGGTTCCGTTGGGCAGTTTTACGGCAA 1447
Db   1448 ATCGTGAATGAATGAGCACGCATACCGGGCGTGCAGTTTGGTAT 1487
```

RESULT 8
US-09-866-379-5
;   Sequence 5, Application US/09866379
;   Patent No. 6855365
; GENERAL INFORMATION:
;   APPLICANT: DIVERSA CORPORATION
;   APPLICANT: SHORT, Jay
;   APPLICANT: KRETZ, Keith
;   APPLICANT: GRAY, Kevin
;   APPLICANT: BARTON, Nelson
;   APPLICANT: GARRETT, James
;   APPLICANT: O'DONOGHUE, Eileen
;   TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
;   FILE REFERENCE: DIVER1370-7
;   CURRENT APPLICATION NUMBER: US/09/866,379
;   CURRENT FILING DATE: 2001-05-24
;   PRIOR APPLICATION NUMBER: US 09/580,515
;   PRIOR FILING DATE: 2000-05-25
;   PRIOR APPLICATION NUMBER: US 09/318,528
;   PRIOR FILING DATE: 1999-05-25
;   PRIOR APPLICATION NUMBER: US 09/291,931
;   PRIOR FILING DATE: 1999-04-13
;   PRIOR APPLICATION NUMBER: US 09/259,214
;   PRIOR FILING DATE: 1999-03-01
;   PRIOR APPLICATION NUMBER: US 08/910,798
;   PRIOR FILING DATE: 1997-08-13
;   NUMBER OF SEQ ID NOS: 10
;   SOFTWARE: PatentIn version 3.1
;
; SEQ ID NO 5
;   LENGTH: 1901
;   TYPE: DNA
;   ORGANISM: Escherichia coli
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1901)
;   OTHER INFORMATION: n is any nucleotide
US-09-866-379-5

Query Match   96.5%;   Score 1262.2;   DB 3;   Length 1901;
Best Local Similarity   98.2%;   Pred. No. 0;   Mismatches   24;   Indels   0;   Gaps   0;
Matches 1276;   Conservative   0;

```
Qy       1 ATGAAAGGGATCTTAATCCCATTTTTATCTCTTCTGATTCGGTTACCCCGCCAAATCTGCA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     188 ATGAAAGGGATCTTAATCCCATTTTTATCTCTTCTGATTCGGTTACCCCGCCAAATCTGCA 247

Qy      61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAAGTGTGGTGATTGTCAGTCGTCATGGT 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     248 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAAGTGTGGTGATTGTCAGTCGTCATGGT 307

Qy     121 GTGCGTGCTCCAACCAAGGCCAACGCAACTGATGCAGGACATGTGCACCCCAGAGGCATGGCCA 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     308 GTGCGTGCTCCAACCAAGGCCAACGCAACTGATGCAGGACATGTGCACCCCAGAGGCATGGCCA 367

Qy     181 ACCTGGCCGGTAAAAACTGGGGTGAGCTGACACCGGCAGACACCGCACGCCGCTATCTC 240
           |||||||||||||||||||||||||||||||||||||||||||||||
Db     368 ACCTGGCCGGTAAAAACTGGGGTTGACTGACACCGGCAGACACCGCACGCCGCTATCTC 427

Qy     241 GGACATTACTGGCGTCAGGTCTGGTAGCCGGACGGATTGCTGCCTAAAAGTGGCTGGCCCG 300
           |||||||||||||||| ||||||||||||||||| ||||||||||||||||||||||||
Db     428 GGACATTACCAACGCCAGGCGTCTGGTAGCCGGATTGCTGGCGAAAAAGGGCTGGCCCG 487

Qy     301 CAGTCTGGTCAGGTCCGGATATATTGTCGATTCGACGAGGCTACCCGTAAAACAGGCGGAA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     488 CAGTCTGGTCAGGTCGGGATTATTGCTGGCGGATGTCGACGAGGCTACCCGTAAAACAGGCGGAA 547

Qy     361 GCCTTCGGCCCCGCCCGCCGGCTGGCACCTGACCTGGCCAATAACCGTACATAACCGTACG 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     548 GCCTTCGGCCTCGCGCCGGGCTGGCACCTGACCTGTGCAATAACCGTACATAACCGTACG 607

Qy     421 TCCAGTCCCGGATCCGGTTATTTAATTCCTCTAAAAACTGGCGTTGCCAACTGGATAACGCG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     608 TCCAGTCCCGGATCCGGTTATTTAATTCCTCTAAAAACTGGGCGTTTGCCAACTGGATAACGCG 667

Qy     481 AACGTGACTGACGCGATCCTTCGGAGGGCAGGGCAGGGGTCAATTGCTGACTTTACCGGGCAT 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     668 AACGTGACTGACGCGATCCTCAGCCAGGGCAGGGGTCAATTGCTGACTTTACCGGGCAT 727

Qy     541 TATCAAAACGGCGGTTTCGGCGGAACTGGAACCGGCTGCTTAATTCTTCCGGCAATCAAACTTGTGC 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     728 CGGCAAAACGGGCGTTTCGGCGGAACTGGCGTGCTTAATTTTCCGCAATCAAACTTGTGC 787

Qy     601 CTTAAACGTGAGAAACAGGACGAAAGCGTTCATTCAAGTTCATTACACTGGAACTC 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     788 CTTAAACGTGAGAAAACAGGAGCGAAAGCGTTCATTAACCGCAGGCAATACCATCGGAACTC 847

Qy     661 AAGGTGAGCGCCGCACTGTGTCTCATTAACCGGTGCGGTGCGGTAACCTCCGCACGCAAAATGCTGACG 720
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db     848 AAGGTGAGCGCGCCGACAAGTCTCATTAACCGGTGCGGTAAGCTCCGCATCAAATGCTGACG 907

Qy     721 GAGATATTTCTCCCTGCAACAAGCACAGGGGAATGCCCGAGCCCGGGGGTGGGGAAGGATCACC 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     908 GAGATATTTCCCTGCAACACGACAGGGGATGCCCGAGCCCGGGGGTGGGGAAGGATCACC 967

Qy     781 GATTCACACAGTGGAACACCTTGCTAAGTTTGCATAACGGCAATTGATTGCTACGA 840
           |||||||||||||||||||||||||||||||||||||||||||||
Db     968 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCCAATTTATTGTTCTACA 1027

Qy     841 CGCACGCCCAGAGGTTGCCCCGCGACCCCGCCCACCCCCGTTATTGAGATTTGATCAAGACAGCG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1028 CGCACGCCAGAGGTTGCCCGCGACCCCGCCCACCCCCGTTATTAGATTTGATCAAGACAGCG 1087

Qy     901 TTGACCCCCCATCCACGCAAAAAACAGGCGTATGGGTGTGACATTACCCACTTCAGTGCTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1088 TTGACCCCCCATCCACGCAAAAACAGGCGTATGGGTGTGACATTACCCACTTCAGTGCTG 1147

Qy     961 TTTATCGCCGGACACGGATACTAATCTGGCAAATCTCGGGGGCACTGGAGCTCAACTGG 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1148 TTTATCGCCGGACACGGATACTAATCTGGCAAATCTCGGGGGCACTGGAGCTCAACTGG 1207

Qy    1021 ACGCTTCCCGCTGCCGGATAACACGCCGCGCAGGTGGTGAACTGGTGTGTTTGAACGCTGG 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1208 ACGCTTCCCGCTGCCGGATAACACGCCGCGCAGGTGGTGAACTGGTGTGTTTGAACGCTGG 1267

Qy    1081 CGTCGGCGCTAAGCGATAACAGACGACGGGGATAACACCGCGCCAGGTGGTCTTCCAGGTTT TACAG 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1268 CGTCGGCGGCTAAGCGATAACAGACCGCCAGTGGGATTTCGGCTGGTTCGCGTTTCGTTCGTCCAGACTTTACAG 1327

Qy    1141 CAGATGCCTGATAAAAACGCCCGTTGCATTAAAATACGGCCGCCCCGGAGAGGAGGTGAAACTGACC 1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1328 CAGATGCCTGATAAAAACGCCCGTTGCATTAAAATACGGCCGCCCCGGAGAGGAGGTGAAACTGACC 1387

Qy    1201 CTGGCAGGATGTGAAGAGAGGCGAAATGCCGAGGGCCGCCAGTGTCGTGTTGGTCAGGTTTTTTACGGCAA 1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1388 CTGGCAGGATGTGAAGAGAGGCGAAATGCCGCAGGGCCGCCATGTTGTTCGTTGGTCAGGTTTTTTACGGCAA 1447

Qy    1261 ATCGTGAATGAAGCACCCGATACCGGCCTGCACTTTGAGAT 1300
           ||||||||||||||||||||||||||||||||||||  ||
Db    1448 ATCGTGAATGAAGCACCGGATACCGGCGTGCAGTTTGTTAAT 1487
```

RESULT 9
US-09-866-379-9
; Sequence 9, Application US/09866379
; Patent No. 6853365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-9

Query Match   96.5%;   Score 1262.2;   DB 3;   Length 1901;
Best Local Similarity   98.2%;   Pred. No. 0;   Mismatches   24;   Indels   0;   Gaps   0;
Matches 1276;   Conservative   0;

```
Qy       1 ATGAAAGGGATCTTAATCCCATTTTTATCTCTTCTGATTCGGTTACCCCGCCAAATCTGCA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     188 ATGAAAGGGATCTTAATCCCATTTTTATCTCTTCTGATTCGGTTACCCCGCCAAATCTGCA 247

Qy      61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAAGTGTGGTGATTGTCAGTCGTCATGGT 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     248 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAAGTGTGGTGATTGTCAGTCGTCATGGT 307

Qy     121 GTGCGTGCTCCAACCAAGGCCAACGCAACTGATGCAGGACATGTGCACCCCAGAGGCATGGCCA 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     308 GTGCGTGCTCCAACCAAGGCCAACGCAACTGATGCAGGACATGTGCACCCCAGAGGCATGGCCA 367
```

RESULT 10
US-09-715-477-2
; Sequence 2, Application US/09715477
; Patent No. 6841370
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
; FILE REFERENCE: 19603/4031
; CURRENT APPLICATION NUMBER: US/09/715,477
; CURRENT FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1486
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-715-477-2

Query Match 94.3%; Score 1233.6; DB 3; Length 1486;
Best Local Similarity 97.0%;
Matches 1257; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 1 ATGAAAGCGATCTTAATCCCATTTTATCTCTTGCGTTAACCCGCCAATCTGCA 60
Db 188 ATGAAAGCGATCTTAATCCCATTTTATCTCTTGCGTTAACCCGCCAATCTGCA 247

Qy 61 TTCGCTCAGAGTGAGAGCCCGAGCTTGAAGCTGGAAAGTGTGGTTGAGTTCAGTCGTCATGGT 120
Db 248 TTCGCTCAGAGTGAGAGCCCGAGCTTGAAGCTGGAAAGTGTGGTTGAGTTCAGCGTCATGGT 307

Qy 121 GTGGGTGTCCAACCCAGCCACCGCAACTGATGCAGCGATTGCCCGGACGTGCCCG 180
Db 308 GTGGGTGCCCCAACCAAGGCCACCGCAACTGATGCGCAGGATTGCCCGGACGTGCCCA 367

Qy 181 ACCTTGGGCCGGTAAAAACTGGGTGGAGCTGACACCCGCGCGGTGGTGAGCTAATCGCTATCTC 240
Db 368 ACCTTGGGCCCGCCCCAACCAAGGCCGTATGGTGACGGAGCGTAAATGCCTATCTC 427

Qy 241 GGACATTACTGGGCGTCGGCGATTGCTGATGTGCGACCGGAGGCGGAA 300
Db 428 CAGCCTGGTCGGCGATTATTGCGCAAGGGCTGCCCG 487

Qy 301 CAGTCTGGTCAGGTCGGCGATTATTGCTAATGTGCGACGAGGGCGGAA 360
Db 488 CAGTCTGGTCAGGTCGGCGATTATTGCGCAATAACCGTACATACCCGAGGCGAA 547

Qy 361 GCCTTCGCCGCGCCGGGCACCTGACTGACTGCAATAACCGTACATACCCGAGGCGAA 420
Db 548 GCCTTCGCCGCGCCGGGCTGACTGTGCAATAACCGTACATACCCGAGGCAGATACG 607

Qy 421 TCCAGTCCGATCCGTTATTTAATCCTCTAAAAAACTGCGGATAACCG 480
Db 608 TCCAGTCCGATCCGTTATTTAATCCTCTAAAAAACTGGATAACGCG 667

Qy 481 AACGTGACTGACGCGATCCTCAGCAGGGCGAGGAGGGTCAATTCGCAGATACG 540
Db 668 AACGTGACTGACGCGATCCTCAGCAGGGCGAGGAGGGTCAATTCGCAGATACG 727

Qy 541 TATCAAACGGCGTTTCGGGAACTGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC 600
Db 728 CGGCAAACGGCGTTTCGGGAACTGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC 787

Qy 601 CTTAAACGTGAGAAAACAGGACAGAAAGCTGTTCATTAACGCAGGCAGGCATTACCATGGAACTC 660
Db 788 CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCATTAACGCAGGCAGGCATTACCATGGAACTC 847

Qy 661 AAGGTGAGCGCGACTGTGCTCTCATTAACCGTGCGGTAAGCCTCGCATCAATGCTGACG 720
Db 848 AAGGTGAGCGCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 907

Qy 721 GAGATATTTCTCCTGCAACAAGCAACAGGGAACGCCGAGCACCGGGGGAGGATCACC 780
Db 908 GAGATATTTCTCCTGCAACAGCACAGGGAATGCCGGAGCGCCGGGGGAAGGATCACC 967

Qy 781 GATTCACACCAGTGGACACCTTGCTAAGTTTGCATAACGCGGCCGGGGGTGGGCTAATCGCCTATCTC 840
Db 968 GATTCACACCAGTGGACACCTTGCTAAGTTTGCATAACGCGGCAATTTGTTGCTACAA 1027

Qy 841 CGCACGCGCAGAGGTTGCCCGCGAGCGCCCGCGCACCCGTTATTAGATTGAATCAAGACAGCG 900
Db 1027 GGACATTACTGGGCTCAGGTCTGGTAGGCGACAGCGATTGCTGACGCGGTAATGTGGGCGTGCCCG 900

Qy 901 TTGACGCGCCCATCCACCGCAAAAAAACAGGCGTATGGTGTGACATTACCCCACTTCAGTCCTG 960
Db 1088 TTGACGCGCCCATCCACGCAAAAAACAGGCGTATGGTGTGACATTACCCCACTTCAGTGCCTG 1147

Qy 961 TTTATCGCCGGACACGATACTTAATCTCGGAGCTCAACTGG 1020
Db 1148 TTTATCGCCGGACACGATACTTAATCTCGGAGCTCAACTGG 1207

Qy 1021 ACGCTTCCCGGTCAGCGCGGATAACACGCGCCGCCAGGTGGTGAACTGGATGGG 1080
Db 1208 ACGCTTCCCGGTCAGCGCGGATAACACGCGCCGCCAGGTGGTGAACTGGGATGGG 1267

Qy 1081 CGTCGGCGTAAGGGATAACAGCCAGTGGATTCGCGTGCTGCTTCGGAGACTTTACAG 1140
Db 1268 CGTCGGCGTAAGGGATAACAGCCAGTGGATTCGCGTGCTCTTCCAGACTTTACAG 1327

Qy 1141 CAGATGGCGTGATAAAACGCCGCTGTCATTAAAATAACGGGTGTGAAACTGACC 1200
Db 1328 CAGATGGCGTAAGCGGATAAAAAGCCCGCGGTTCATTAAAATAACTGGGAACGGGGTGAAACTGACC 1387

Qy 1201 CTGGGCAGGAGGATGTGAAGAGCGAAAATGCCCAGGGCGCATGTGTTCGTTGGCAGGTTTTTACGGCAA 1260
Db 1388 CTGGGCAGGATGTGAAGAGCGAAAATGCCCAGGGCGCATGTGTTCGTTGGCAGGTTTTTACGGCAA 1447

Qy 1261 ATCGTGAATGAAGAAGGCACGCAATACCGGCGTCCAGTTTGAGAT 1300
Db 1448 ATCGTGAATGAAGCACACCAATACCCGGCGTCCAGTTTGTAAT 1487

RESULT 11
US-09-540-149A-9
Sequence 9, Application US/0954O149A
; Patent No. 6511699
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2791
; CURRENT APPLICATION NUMBER: US/09/540,149A
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 60/127,032
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1489
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-540-149A-9

Query Match   93.9%; Score 1228; DB 3; Length 1489;
Best Local Similarity 96.5%;
Matches 1255; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

(The remainder of the page consists of a rotated nucleotide sequence alignment between Query (Qy) and Database (Db) sequences, shown as paired rows with flanking coordinate numbers and match bars, e.g.:)

```
Qy    1 ATGAAAGCGATCTTAATCCCATTTTATCTCTTTTTGATTCGTTAACCCCGGCAATCTGCA   60
Db  182 ATGAAACGGATCTTAATCCCATTTTATCTCTTTGATTCGTTAACCCCGGCAATCTGCA   241
```

RESULT 12
US-10-266-041A-9

; Sequence 9, Application US/10266041A
; Patent No. 6974690
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: PHOSPHATASESS WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2298
; CURRENT APPLICATION NUMBER: US/10/266,041A
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: 60/127,032
; PRIOR FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 09/540,149
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 9
; LENGTH: 1489
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-266-041A-9

Query Match   93.9%;   Score 1228;   DB 4;   Length 1489;
Best Local Similarity   96.5%;   Pred. No. 0;
Matches 1255;   Conservative   0;   Mismatches 45;   Indels 0;   Gaps 0;

Qy   1 ATGAAAAGCGGATCTTAATCCATTTTATCTCTTCTGATTCCGTTAACCCGCAATCTGCA   60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   182 ATGAAAAGCGGATCTTAATCCATTTTATCTCTTCTGATTCCGTTAACCCGCAATCTGCA   241

Qy   61 TTCGCTCAAGAGTGAGCCGGAGCTGAAGCTGGAAAAGTGTGGTGAGTCGAGTCGTCTGGT   120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   242 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAAGTGTGGTGAGTCGAGTCGTCATGGT   301

Qy   121 GTGCGTGCTCCAACCAAGGCCAGCAACTGATGCAGGACGCAGCGGCATGGCCA   180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   302 GTGCGTGCCCCAACCAAGGCACGCAACTGATGCAGGACGCGGCAAATGGGCCATGGCCA   361

Qy   181 ACCTGGCGCGGTAAAACTGGGTGAGCTGACTGACACCGCGCGGCTAATGCCCTACTC   240
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   362 ACCTGGCGCGGTAAAACTGGGTGTGGCTGACACACCGCGGGGTGGCTAATCGGCCTACTC   421

Qy   241 GGACATTACTGCGGCGGTCAGCGCGTCTGGGTAGCCGACGGGATTGCTGCGGGCCTGCGCCG   300
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   422 GGACATTACCAACGCCAGCGTCTGGGTGCCGACGGAAAAACAGGGCCTGCCCG   481

Qy   301 CAGTCTGGTCAGGTCGGCGGATTATTGCTGAATGCGAACCGTACCCGTAAAACAGGGCGAA   360
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   482 CAGCCTCGGTCAGGTCGGGATTATTGCTGATGTCGACAGCGGTACCCCGTAAAACAGGGCGAA   541

Qy   361 GCCTTCGCGCCGCCCGGCTGGCACCTGACTGTTGCAAATAACCGTACTACCCAGGCCAGATACG   420
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   542 GCCTTCGCGCCGCCGGGCTGGCACTGACTGTTGCAAATAACCGTACTACCCAGGCCAGATACG   601

Qy   421 TCCAGTCCCGATCCGGTTATTTAATTCCTCTAAAAAACTGGCCGTTTGCCAACTGGATAACGCG   480
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   602 TCCAGTCCCGATCCGGTTATTTAATTCCTCTAAAAAACTGGCCGTTTGCCAACTGGATAACGCG   661

Qy   481 AACGTGACTGACGCGGATCCTCGAGAGGGGCGCAGGAGGTCAATTGCTGACTTTACCGGGGCAT   540
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   662 AACGTGACTGACGCGGATCCTCAGCGAGGGGCGCAGGAGGTCAATTGCTGACTTTACCGGGGCAT   721

Db   1201 CTGGGCAGGATGTGAAAATGCCGCAGGGGCATGTGTTCGTTCGGTTGGCCGGTTTTACCGGCAA   1260

Qy   1201 ATCGTGAATGAAGCGCCATACCGGCGTGCAGTTTGTGAT   1300

Db   1382 CTGGGCAGGATGTGGAAAATGCCGCAGGGGCATGTGTTCGTTCGGTTGGCCGGTTTTACCGGCAA   1441

Qy   1382 CTGGGCAGGATGTGGAAAATGCCGCAGGGGCATGTGTTCGTTCGGTTGGCCGGTTTTACCGGCAA   1441

Qy   1442 ATCGTGAATGAAGCGCCATACCGGCGTGCAGTTTGTGAT   1481

Db   1322 CAGATGCCTGATAAAACCGCGCTGTCATTAAAATACAGCGGTTTCAGGTTTCAGGTTTCGTCATGGTTCATTGGT   1381

RESULT 13
US-09-715-477-4

; Sequence 4, Application US/09715477
; Patent No. 6841370
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
; FILE REFERENCE: 19603/4031
; CURRENT APPLICATION NUMBER: US/09/715,477
; CURRENT FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 4
; LENGTH: 1486
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-715-477-4

```
Query Match      93.5%;                        Score 1222.4;   DB 3;   Length 1486;
Best Local Similarity   96.5%;   Pred. No. 0;   Mismatches 46;   Indels 0;   Gaps 0;
Matches 1250;   Conservative   0;   Mismatches   46;   Indels   0;   Gaps   0;

Qy      1  ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTCGATTCCGTTAACCCCGCCAATCTGCA       60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188  ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTCGATTCCGTTAACCCCGCCAATCTGCA      247

Qy     61  TTCCGCTCAGAGTGAGCCGGAGCTGAGAAGTGGTGTGATTGTCAGTCGTCGTCATGGT      120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248  TTCCGCTCAGAGTGAGCTGGAGCTGAGAAGTGTGGTGATTGTCAGTCGTCATGGT      307

Qy    121  GTGCGTGCTTCCAACCAAAGGGCCAAACTGATGCCAGACGGCCATGGGGCCA      180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308  GTGCGTGCTTCCCAACCAAAGGGCCACGAACTGCATGCAGGATGTCACCCCCAGACGGCCA      367

Qy    181  ACCTGGCCGGGTAAAACTGGGTGAGCTGACACCGCGCGGGTGGGTGGGGTAATGCCTATCTC      240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    368  ACCTGGCCGGGTAAAACTGGGTTGGCTGACACCGCGCGGGGTGGGTAATGCCTATCTC      427

Qy    241  GGACATTACTGGGCGTCAGCGTCTCGTGTTGCTGCCTAAAATGTGGCTGCCCG      300
            ||||||||||||||||||||||||||||||||||||||||||||
Db    428  GGACATTACCAACGCCAGCCGTTCTGGTGGCCGGACGGATTGCTGGCGAAAAGGCTGCCCG      487

Qy    301  CAGTCTGGTCAGGTCGCGGATTATTGCTGTATGATCGACGAGGCTACCCGGTAAAAACAGG      360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488  CAGCGCTGGTCAGGTCGCGGATTATTGTCGACGATGTCGACGAGGCTACCCGGTAAAAACAGG      547

Qy    361  GCCTTCGCCCGCGGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG      420
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db    548  GCCTTCGCCCGCGGGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG      607

Qy    421  TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCCGTTTGCCAACTGGGATAACGCG      480
            ||||||||||||||||||||||||||||||||||||||||||||
Db    668  AACGTGACTAGCGCATCCTCAGCAGGGTCAATTGCTGACTTTTACCGGGCAT      727

Qy    481  AACGTGACTAGCGCATCCTCGAGAGGGCAGGAGGGTCAATTGCTGACTTTTACCGGGCAT      540
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    668  AACGTGACTAGCGCATCCTCAGCAGGGTCAATTGCTGACTTTTACCGGGCAT      727

Qy    541  TATCAAAACGGCGTTTGGGAACTGGAACGGCGTGCTTAATTTTCCGAATCATCAAACTTGTGC      600
            ||||||||||||||||||||||||||||||||||||||||||
Db    728  CGGCAAAACGGCGTTTGCGGAACTGGAACGGCGTGCTTAATTTTCCGAATCATCAAACTTGAC      787

Qy    601  CTTAAAACGTGAGAAACAGGACGGAAAGCTGTTCATTAACGACGAGGCAGCATTACCATTCGGAACTC      660
            ||||||||||||||||||||||||||||||||||||||||||
Db    728  CGGCAAAACGGCGTTTGCGGAACTGGAACGGCGTGCTTAATTTTCCGAATCATCAAACTTGAC      787

Qy    661  AAGGTGAGCGCCGACTGTGTTGTCTCATTAACCCGGTGCGGTAGCCCTCGCCATCAATGCTGACG      720
            ||||||||||||||||||||||||||||||||||||||||
Db    848  AAGGTGAGCGCCGACAAATGTTTCATTAACCCGGTGCGGGATCACCG      907

Qy    721  GAGATATTTTCTTCCTCGAACAAGCAACAGGGAATGCCGGAGCCCGGAGGAGGGGAGGGATCACC      780
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db    848  AAGGTGAGCGCCGACAAATGTTTCATTAACCCGGTGCGGGATCACCG      907

Qy    781  GATTCACACCAGTGGAACACCTTGCTTGCTAAGTTTGCATAACCGGCAATTTGATTTTGCTACAA      840
            |||||||||||||||||||||||||||||||||||||||||||||||
Db    968  GATTCACACCAGTGGAACACCTTGCTTGCTAAGTTTGCATAACCGGCAATTTTATTTACTACGA     1027

Qy    841  CGCACGCCCAGAGGTTGCCCGCGCCAGCCCGCCACCCCCGTTATTGGATTTGATCAAGACAGCG      900
            ||||||||||||||||||||||||||||||||||||||||||||||
Db    968  GATTCACACCAGTGGAACACCTTGCTTGCTAAGTTTGCATAACCGGCAATTTTATTTACTACGA     1027

Qy    901  TTGACGGCGCCATCCACCGCAAAAAAACAGGCGTAGGTGGTGTGACATTACCCACTTCAGTGCTG      960
            ||||||||||||||||||||||||||||||||||||||||||
Db   1088  TTGACGGCCCCATCCACCGCAAAAACAGGCGTATGGGGTGTGACATTACCCACTTCAGTGCTG     1147

Qy    961  TTTATCGCCGGACACGATACTAATCTCGGGCAAAATCTCGGGCGCGGCCCACTGGAGCTCAACTGG     1020
            ||||||||||||||||||||||||||||||||||||||||||
Db   1088  TTGACGGCCCCATCCACCGCAAAAACAGGCGTATGGGGTGTGACATTACCCACTTCAGTGCTG     1147

Qy   1028  CGCACGCCCAGAGGTTGCCCGCCAGTCCGCCGCCACCCCCGTTATTGGATTTGATCAAGACAGCG     1087

Qy   1021  ACGCTTCCCGCGTCAGCCGGATAACGACGCCCAGGTTGGTGGTGAAACTGGCGTGGTTGTTGAACCGTGG     1080
            ||||||||||||||||||||||||||||||||||||||||||
Db   1208  ACGCTTCCCAGGTCAGCCGGATAACGACGCCCAGGTTGGTGGTGAAACTGGCGTGGTTGTTGAACCGTGG     1267

Qy   1081  CGTCCGGCTAAGCCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTGTCTTCCAGACTTTACAG     1140
            ||||||||||||||||||||||||||||||||||||||||||
Db   1268  CGTCCGGCTAAGCCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTGTCTTCCAGACTTTACAG     1327

Qy   1141  CAGATGCCGTGATAAAAGCCCGCTGTCATTAAAATACGGCCCCGGGAGGAGGGTGAAACTGACC     1200
            ||||||||||||||||||||||||||||||||||||||||||
Db   1328  CAGATGCCGTGATAAAAGCCCGCTATCATTAAAATACCGCCCCGGGAGGAGGGTGAAACTGACC     1387

Qy   1201  CTGGCAGGATGTGAAGAGCGGAAATGGGCAGGGCCAGGGCATGTTGCCAGGTTTTTACGGCAA     1260
            ||||||||||||||||||||||||||||||||||||||||||
Db   1388  CTGGCAGGATGTGAAGAGCGGAAATGGGCAGGGCATGTTGGTTGGCCAGGTTTTTACGGCAA     1447

Qy   1261  ATCGTGAATGAAGCACGCATACCGGCGGTGCAGTTTG     1296
            |||||||||||||||||||||||||||||||||||||
Db   1448  ATCGTGAATGAAGCGCGCCATACCGGCGGTGCAGTTTG     1483
```

```
Query Match      88.1%;                        Score 1153;   DB 2;   Length 1272;
Best Local Similarity   94.1%;   Pred. No. 0;   Mismatches 26;   Indels 51;   Gaps 1;
Matches 1229;   Conservative   0;   Mismatches   26;   Indels   51;   Gaps   1;

Qy      1  ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTCGATTCCGTTAACCCCGCCAATCTGCA       60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTCGATTCCGTTAACCCCGCCAATCTGCA       60
```

```
Qy    61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCGTCATGGT 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCGTCATGGT 120

Qy   121 GTGCGTGCTCCAACCAAGCCACCAACTGATCGCAGATGTCACCCCAGACGCATGGGCCA 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTGCGTGCTCCAACCAAGCCACCAACTGATCGCAGATGTCACCCCAGACGCATGGGCCA 180

Qy   181 ACCTGGCGCGGTAAAACTGGGTGAGCTGAGCGTAATCGGCTTACTCTC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ACCTGGCGCGGTAAAACTGGGTGAGCTGAGCGTAATCGGCTTACTCTC 240

Qy   241 GGACATTACTGGGGTCAGCGTCTGGGTAGCCGACGGATTGCTGCGCCCG 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGACATTACCAACGCCAGCGTCTCGGTAGCCGACGGATTGCTGCGCCCG 300

Qy   301 CAGTCTGGTCAGGTCGCGATTATTGCTGCAGCGTACCCGTAAAACAGG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGTCTGGTCAGGTCGCGATTATTGCTGCAGCGTACCCGTAAAACAGG 356

Qy   357 ---------------------------------CCAGGCAGATACG 369
                                          ||||||||||||
Db   357 ---------------------------------CCAGGCAGATACG 369

Qy   361 GCCTTCGCCGCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCCTTCGCCGCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420

Qy   421 TCCAGTCCGATCCGTTATTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAAGGCG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCCAGTCCGATCCGTTATTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAAGGCG 480

Qy   481 AACGTGACTGACCGATCCTCAGCAGGGCAGGAGGGTCAATTCGTACTTTACCGGGGCAT 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AACGTGACTGACCGATCCTCAGCAGGGCAGGAGGGTCAATTCGTACTTTACCGGGGCAT 489

Qy   541 TATCAAACGGCGTTTCGCGGAACTGGGAACGGGGTGCTTAATTTTCCGCAATCAAAACTTGTGC 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CGGCAAACGGCGTTTCGCGGAACTGGGAACGGGGTGCTTAATTTTCCGCAATCAAAACTTGTGC 549

Qy   601 CTTAAAGTGAGAAACAGGACAGAAAGCTGTGCTCATTAACGGCAGCCATTACCATCGGAACTC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTTAAAGTGAGAGAAAACAGGACGAAAGCTGTTCATTAACGGCAGCCATTACCATCGGAACTC 660

Qy   661 AAGGTGAGCGCCGACTGTGTCTCATTAACGGCGGTGGGTTAAGCCTTCGCATCATCAATGCTGACG 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGGGTTAAGCCTTCGCATCATCAATGCTGACG 669

Qy   721 GAGATATTTCTCCTGCAACAGCCACAGGGAATGCGGCGACCGGCGGGGTGGGGAAGGATCACC 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GAGATATTTCTCTGCAACAGCAGCACAGGGAATCCGGGCGGGGTGGGGAAGGATCACC 729

Qy   781 GATTCACACCAGTGGGAACACCTTGCTAAGTGTTGGCATAAGCGGCGCAATTGATTTGCTACAA 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GATTCACACCAGTGGGAACACCTTGCTAAGTTTGCATAAGCGGCGCAATTGATTTTGCTACAA 840

Qy   841 CGCACGCGCAGAGGTTGCCCGCAGCCGGGGCCACCCCGGTTATTGAGATTGGATCAAGACAGCG 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CGCACGCGCAGAGGTTGCCCGCAGCCGGGGCCACCCCGGTTATTGAGATTTGGATCATGGCAGCG 849

Qy   901 TTGACGCGCCCATCCACGCGCAAAAACAGCGGCGTATGGTGACATTACCCAATTCAGTGCTG 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTGACGCGCCCATCCACGCGCAAAATCTCGGCGTATGGTGACATTACCCAACTTCAGTACTG 909

Qy   961 TTTATCGCCGGACACGATACTAATCTCGGCGGCCACTGGAGTCAACTGG 1020
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db   910 TTTATTTGCGGGACACGATACTAATCTCGGCGGCCACTGGAGTCAACTGG 969

Qy  1021 ACGCTTCCCGGTCAGCGGTCCGGACACGATACTCAGTTGGTGTTTTGAACGCTTACAG 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   970 ACGCTTCCCGGTCAGCGGTAACACGGCCGGAATAACAGCCAGTGGATTCAGGTTTCGCTGG 1029

Qy  1081 CGTCGGCTAAGCGGATAACAGCCAGTGGATTCAGGTTTCGCTGGTTCTTCCAGACTTTACAG 1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1030 CGTCGGCTAAGCGGATAACAGCCAGTGGATTCAGGTTTCGCTGGTTCTTCCAGACTTTACAG 1089
```

```
Query Match       5.1%;  Score 66.2; DB 3;  Length 1266;
Best Local Similarity  49.2%;  Pred. No. 9.4e-11;
Matches  204; Conservative  0; Mismatches  208; Indels  3; Gaps  1;

Qy  1141 CAGATGCGTGATGATAAAAACCGGCGTCTGCTACTTCATTAAAATACGCCCGCCGGGAGGGAGAGGTGAAAACTGACC 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1090 CAGATGCGTGATGATAAAAACCGGCGTCTGCTACTTCATTAAAATACGCCCGCCGGGAGGGAGAGGTGAAAACTGACC 1149

Qy  1201 CTGGCCAGGATGTGAAGAGGGAAGAGGCATGTGTTCGGTTCGGTTGGCAGGTTTTTACGCAA 1260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1150 CTGGCCAGGATGTGAAGAGGGAAGAGGCATGTGTTCGGTTCGGTTGGCAGGTTTTTACGCAA 1209

Qy  1261 ATCGTGAATGAAGCACGCAATACCCGGCGTCGCAGTTTGAGATCTCATC 1306
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1210 ATCGTGAATGAAGCACGCAATACCCGGCGTCGCAGTTTGAGATCTCATC 1255
```

```
Qy    86 AGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGTGTCGTGCTCC---AACCAAGGCCA 142
         |||||||||||||||||||||||||||||||||||||||||||||   |||||||||||
Db   113 AGCTGCAGCAGGTTGCCTGATCATGAGCCGTGCCGTGCCCGTTGGCCAATAACG 172

Qy   143 CGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCAACCTGGCGCGGTAAAAACTGGGTG 202
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   173 GCAGCGTGCTGGAACAGTCCACCGCCAAGGCGTCTGGGCGGATGTGCCCGGCGCGCGGCC 232

Qy   203 AGCTGACACCGCGCGGGTGGGTGAGCTAATCGCCTATCTCGGACATTACTGGCGTCAGCGTC 262
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   233 AGCTGACCACCAAAAGGCGCGGTGTTGCTCGGAGGTGTATATGGGCGACCACTATATGCGCGAATGGC 292

Qy   263 TGGTAGCCGACGGATTGCTGCGCCCCCAGTCTAAAATGGGCTGGGCGCCCCAGTCTGGTCAGGTTCGGGGCGATTA 322
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   293 TGGCCGCAGCAGGAAGCTGGTTGACCAGGCGCGAGTGCCCGCCGGGAAAACGGGGTTTATGCCGT 352

Qy   323 TTGCTGATGTGTCGACGGACGGGTACCCGGTAAAAACAGGCGGGAAGCCTTCGCCGCCCGGTCGGTCAC 382
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   353 ACCGTAACAGCCTCCGCCCACGCTGCCGCAGTTCTTTATCACCGGCGGCGGTGTCC 412

Qy   383 CTGACTGTGCAATAACCGGTACATACCCAGGCAGATAAGTCCAGTCCCGGATCCGTTATTTA 442
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   413 CGGGCTGCGGGCATCCCGGTGCATCACCAGCCTCGAGATGGGCACCAACCCCGACCCTTCA 472

Qy   443 ATCCTCTAAAAAAACTGGCGTTTGCCAACTGGATAACGCTGACTGACCGAT 497
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   473 ACCCGGTTATTACCGACGACTCCCCGCGGGTTTGCGGAAGCCCTGCAGGGCGAT 527
```

THIS PAGE LEFT BLANK

OM protein - protein search, using sw model

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

Run on: June 13, 2006, 10:45:37 ; Search time 143.303 Seconds
         (without alignments)
         1422.265 Million cell updates/sec

Title: US-10-601-319-2
Perfect score: 2302
Sequence: 1 MKAILIPFLSLLIPLTPQSA........IVNEARIPACSLRSHHHHHH 440

Scoring table: BLOSUM62
         Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 46321485 residues

Total number of hits satisfying chosen parameters: 2097797

Post-processing: Minimum Match 0%
         Maximum Match 100%
         Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Database :   Published Applications AA Main:*
   1:   /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
   2:   /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
   3:   /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
   4:   /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
   5:   /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
   6:   /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep:*

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2302	100.0	440	3	US-09-777-566A-2	Sequence 2, Appli
2	2302	100.0	440	4	US-09-866-379-2	Sequence 2, Appli
3	2302	100.0	440	4	US-10-034-985-2	Sequence 2, Appli
4	2302	100.0	440	4	US-10-430-356-2	Sequence 2, Appli
5	2302	100.0	440	5	US-10-601-319-2	Sequence 2, Appli
6	2302	100.0	440	5	US-10-933-115-2	Sequence 2, Appli
7	2302	97.1	432	4	US-09-866-379-8	Sequence 8, Appli
8	2235	97.1	432	4	US-10-156-660-4	Sequence 4, Appli
9	2235	97.1	432	4	US-09-866-379-4	Sequence 4, Appli
10	2235	97.1	432	4	US-10-282-122A-43351	Sequence 43351, A
11	2235	97.1	432	4	US-10-601-319-8	Sequence 8, Appli
12	2235	97.1	432	5	US-10-933-115-8	Sequence 8, Appli
13	2235	97.1	432	5	US-10-472-317-41	Sequence 41, Appli
14	2230	96.9	432	6	US-11-056-354-4	Sequence 4, Appli
15	2221	96.5	432	6	US-11-018-709-1	Sequence 1, Appli
16	2221	96.5	432	6	US-10-284-962-3	Sequence 3, Appli
17	2221	96.5	433	6	US-10-266-041-1	Sequence 1, Appli
18	2210	96.0	432	6	US-11-018-709-3	Sequence 3, Appli
19	2197	95.4	432	6	US-10-284-962-5	Sequence 5, Appli
20	2188	95.0	436	4	US-10-156-660-2	Sequence 2, Appli
21	2171	94.3	432	4	US-10-601-319-10	Sequence 10, Appli
22	2171	94.3	432	5	US-10-933-115-10	Sequence 10, Appli
23	2171	94.3	432	6	US-11-056-354-2	Sequence 2, Appli
24	2165	93.0	432	6	US-10-021-723A-16	Sequence 16, Appli
25	2141	93.0	430	4	US-09-866-379-10	Sequence 10, Appli
26	2081.5	90.4	437	6	US-11-074-522-12	Sequence 12, Appli
27	2070	89.9	412	4	US-10-334-672-1	Sequence 1, Appli
28	2070	89.9	412	4	US-10-334-671-1	Sequence 1, Appli
29	2066	89.7	410	4	US-10-021-723A-13	Sequence 13, Appli
30	2066	89.7	410	4	US-10-021-723A-15	Sequence 15, Appli
31	1203	52.3	261	4	US-10-450-763-56057	Sequence 56057, A
32	960.5	52.3	261	5	US-10-021-723A-12	Sequence 12, Appli
33	943.5	41.0	441	4	US-10-282-122A-77792	Sequence 77792, A
34	935.5	41.7	441	4	US-10-021-723A-2	Sequence 2, Appli
35	932.5	40.6	441	4	US-10-021-723A-4	Sequence 4, Appli
36	760	40.5	420	4	US-10-021-723A-8	Sequence 8, Appli
37	657	33.0	476	4	US-10-021-723A-10	Sequence 10, Appli
38	611.5	28.5	144	5	US-10-450-763-54615	Sequence 54615, A
39	476.5	26.7	318	4	US-10-021-723A-14	Sequence 14, Appli
40	472.5	20.7	409	4	US-10-021-723A-6	Sequence 6, Appli
41	412.5	17.9	441	6	US-11-098-686-10683	Sequence 10683, A
42	326	14.2	118	4	US-10-450-763-56055	Sequence 56055, A
43	195	8.5	426	4	US-10-257-174-44	Sequence 44, Appli
44	195	8.5	426	4	US-10-343-357-6	Sequence 6, Appli
45	195	8.5	426	5	US-10-450-763-53707	Sequence 53707, A

ALIGNMENTS

RESULT 1
US-09-777-566A-2

; Sequence 2, Application US/09777566A
; Patent No. US20010055788A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-6
; CURRENT APPLICATION NUMBER: US/09/777,566A
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n is any nucleotide
US-09-777-566A-2

Query Match   100.0%;   Score 2302;   DB 3;   Length 440;
   Best Local Similarity   100.0%;
   Matches   440;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  241 EIFLLQQAQGMPEPGPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  421 IVNEARIPACSLRSHHHHHH 440
        |||||||||||||||||||||
Db  421 IVNEARIPACSLRSHHHHHH 440
```

RESULT 2
US-09-866-379-2
; Sequence 2, Application US/09866379
; Patent No. US2002013675A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-866-379-2

Query Match   100.0%; Score 2302; DB 3; Length 440;
Best Local Similarity   100.0%; Pred. No. 2.2e-193;
Matches 440; Conservative   0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
```

```
Qy  241 EIFLLQQAQGMPEPGPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
```

RESULT 3
US-10-034-985-2
; Sequence 2, Application US/10034985
; Publication No. US20030049815A1
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/034,985
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; SOFTWARE: FastSEQ for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-034-985-2

Query Match   100.0%; Score 2302; DB 4; Length 440;
Best Local Similarity   100.0%; Pred. No. 2.2e-193;
Matches 440; Conservative   0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  241 EIFLLQQAQGMPEPGPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 4
US-10-430-356-2
; Sequence 2, Application US/10430356
; Publication No. US20030232041A1
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/430,356
; CURRENT FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-430-356-2

Query Match   100.0%; Score 2302; DB 4; Length 440;
Best Local Similarity 100.0%; Pred. No. 2.2e-193;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy   121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy   241 EIFLLQQAQGMPEPGWGKRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 EIFLLQQAQGMPEPGWGKRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy   421 IVNEARIPACSLRSHHHHHH 440
          ||||||||||||||||||||
Db   421 IVNEARIPACSLRSHHHHHH 440

RESULT 5
US-10-601-319-2
; Sequence 2, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
;   AND USING THEM
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-601-319-2

Query Match   100.0%; Score 2302; DB 4; Length 440;
Best Local Similarity 100.0%; Pred. No. 2.2e-193;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy   121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy   241 EIFLLQQAQGMPEPGWGKRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 EIFLLQQAQGMPEPGWGKRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy   421 IVNEARIPACSLRSHHHHHH 440
          ||||||||||||||||||||
Db   421 IVNEARIPACSLRSHHHHHH 440

RESULT 6
US-10-933-115-2
```
; Sequence 2, Application US/10933115
; Publication No. US20050281792A1
; GENERAL INFORMATION:
;   APPLICANT: Short, Jay M.
;   APPLICANT: Kretz, Keith A.
;   APPLICANT: Gray, Kevin A.
;   APPLICANT: Barton, Nelson Robert
;   APPLICANT: Garrett, James B.
;   APPLICANT: O' Donoghue, Eileen
;   APPLICANT: Mathur, Eric J.
;   TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
;   TITLE OF INVENTION: THEREOF
;   FILE REFERENCE: 564462001822
;   CURRENT APPLICATION NUMBER: US/10/933,115
;   CURRENT FILING DATE: 2004-09-01
;   PRIOR APPLICATION NUMBER: US/09/866,379
;   PRIOR FILING DATE: 2001-05-24
;   PRIOR APPLICATION NUMBER: US 09/580,515
;   PRIOR FILING DATE: 2000-05-25
;   PRIOR APPLICATION NUMBER: US 09/318,528
;   PRIOR FILING DATE: 1999-05-25
;   PRIOR APPLICATION NUMBER: US 09/291,931
;   PRIOR FILING DATE: 1999-04-13
;   PRIOR APPLICATION NUMBER: US 09/259,214
;   PRIOR FILING DATE: 1999-03-01
;   PRIOR APPLICATION NUMBER: US 08/910,798
;   PRIOR FILING DATE: 1997-08-13
;   NUMBER OF SEQ ID NOS: 10
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 440
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-10-933-115-2

Query Match        100.0%;  Score 2302;  DB 5;  Length 440;
Best Local Similarity  100.0%;  Pred. No. 2.2e-193;
Matches  440;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy  421 IVNEARIPACSLRSHHHHHH 440
          ||||||||||||||||||||
Db  421 IVNEARIPACSLRSHHHHHH 440
```

RESULT 7
US-09-866-379-8
```
; Sequence 8, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
;   APPLICANT: DIVERSA CORPORATION
;   APPLICANT: SHORT, Jay
;   APPLICANT: KRETZ, Keith
;   APPLICANT: GRAY, Kevin
;   APPLICANT: BARTON, Nelson
;   APPLICANT: GARRETT, James
;   APPLICANT: O'DONOGHUE, Eileen
;   TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
;   FILE REFERENCE: DIVER1370-7
;   CURRENT APPLICATION NUMBER: US/09/866,379
;   CURRENT FILING DATE: 2001-05-24
;   PRIOR APPLICATION NUMBER: US 09/580,515
;   PRIOR FILING DATE: 2000-05-25
;   PRIOR APPLICATION NUMBER: US 09/318,528
;   PRIOR FILING DATE: 1999-05-25
;   PRIOR APPLICATION NUMBER: US 09/291,931
;   PRIOR FILING DATE: 1999-04-13
;   PRIOR APPLICATION NUMBER: US 09/259,214
;   PRIOR FILING DATE: 1999-03-01
;   PRIOR APPLICATION NUMBER: US 08/910,798
;   PRIOR FILING DATE: 1997-08-13
;   NUMBER OF SEQ ID NOS: 10
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;   LENGTH: 432
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-09-866-379-8

Query Match         97.1%;  Score 2235;  DB 3;  Length 432;
Best Local Similarity   99.5%;  Pred. No. 1.7e-187;
Matches  430;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Db  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy  421 IVNEARIPACSL 432
          ||||||||||||
Db  421 IVNEARIPACSL 432
```

RESULT 8

US-10-156-660-4
; Sequence 4, Application US/10156660
; Publication No. US20030109958A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson R.
; APPLICANT: Garrett, James B.
; APPLICANT: O'Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM
; TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 09010-029007
; CURRENT APPLICATION NUMBER: US/10/156,660
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
;
; SEQ ID NO 4
; TYPE: PRT
; LENGTH: 432
; ORGANISM: Escherichia coli
US-10-156-660-4

Query Match   97.1%; Score 2235; DB 4; Length 432;
Best Local Similarity   99.5%; Pred. No. 1.7e-187;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60

Qy   61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120

Qy  121  APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180

Qy  181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240

Qy  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA  300

Qy  301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNIANLGGALELNWTLPGQPDNTPPGGELVFERW  360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNIANLGGALELNWTLPGQPDNTPPGGELVFERW  360

Qy  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420

Qy  421  IVNEARIPACSL  432
         ||||||||||||
Db  421  IVNEARIPACSL  432
```

RESULT 9
US-10-282-122A-43351
; Sequence 43351, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 78614
; Remaining Prior Application data removed - See File Wrapper or PALM.
; SOFTWARE: Patentin version 3.1
;
; SEQ ID NO 43351
; TYPE: PRT
; LENGTH: 432
; ORGANISM: Escherichia coli
US-10-282-122A-43351

Query Match   97.1%; Score 2235; DB 4; Length 432;
Best Local Similarity   99.5%; Pred. No. 1.7e-187;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60

Qy   61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120

Qy  121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180

Qy  181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240

Qy  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA  300
```

RESULT 10
US-10-601-319-8
; Sequence 8, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
;   APPLICANT: Short, Jay M.
;   APPLICANT: Kretz, Keith A.
;   APPLICANT: Gray, Kevin A.
;   APPLICANT: Barton, Nelson Robert
;   APPLICANT: Garrett, James B.
;   APPLICANT: O' Donoghue, Eileen
;   APPLICANT: Mathur, Eric J
;   TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
;   TITLE OF INVENTION: AND USING THEM
;   FILE REFERENCE: 09010-029011
;   CURRENT APPLICATION NUMBER: US/10/601,319
;   CURRENT FILING DATE: 2003-06-20
;   PRIOR APPLICATION NUMBER: US 09/291,931
;   PRIOR FILING DATE: 1999-04-13
;   PRIOR APPLICATION NUMBER: US 09/259,214
;   PRIOR FILING DATE: 1999-03-01
;   PRIOR APPLICATION NUMBER: US 08/910,798
;   PRIOR FILING DATE: 1997-08-13
;   PRIOR APPLICATION NUMBER: US 09/866,379
;   PRIOR FILING DATE: 2001-05-24
;   PRIOR APPLICATION NUMBER: US 09/580,515
;   PRIOR FILING DATE: 2000-05-25
;   PRIOR APPLICATION NUMBER: US 09/318,528
;   PRIOR FILING DATE: 1999-05-25
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   NUMBER OF SEQ ID NOS: 10
;   SEQ ID NO 8
;   LENGTH: 432
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-10-601-319-8

  Query Match   97.1%; Score 2235; DB 4; Length 432;
  Best Local Similarity   99.5%; Pred. No. 1.7e-187;
  Matches   430; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

```
Qy    1 MKAILIPFLSLLILPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLILPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy  421 IVNEARIPACSL 432
          ||||||||||||
Db  421 IVNEARIPACSL 432
```

RESULT 11
US-10-472-317-41
; Sequence 41, Application US/10472317
; Publication No. US20040185562A1
; GENERAL INFORMATION:
;   APPLICANT: Cargill, Incorporated
;   TITLE OF INVENTION: Myo-Inositol Oxygenases
;   FILE REFERENCE: 10829/003US1
;   CURRENT APPLICATION NUMBER: US/10/472,317
;   CURRENT FILING DATE: 2003-09-17
;   PRIOR APPLICATION NUMBER: PCT/US02/08404
;   PRIOR FILING DATE: 2002-03-19
;   PRIOR APPLICATION NUMBER: US 60/277,148
;   PRIOR FILING DATE: 2001-03-19
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   NUMBER OF SEQ ID NOS: 72
;   SEQ ID NO 41
;   LENGTH: 432
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-10-472-317-41

  Query Match   97.1%; Score 2235; DB 4; Length 432;
  Best Local Similarity   99.5%; Pred. No. 1.7e-187;
  Matches   430; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

```
Qy    1 MKAILIPFLSLLILPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLILPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy  421 IVNEARIPACSL 432
          ||||||||||||
Db  421 IVNEARIPACSL 432
```

RESULT 12
US-10-933-115-8
; Sequence 8, Application US/10933115
; Publication No. US20050281792A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 564462001822
; CURRENT APPLICATION NUMBER: US/10/933,115
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO 8
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
; SOFTWARE: FastSEQ for Windows Version 4.0

US-10-933-115-8

Query Match 97.1%; Score 2235; DB 5; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.7e-187;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60

Qy     61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120

Qy    121  APAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  APAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180

Qy    181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240

Qy    241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300

Qy    301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360

Qy    361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420

Qy    421  IVNEARIPACSL  432
           ||||||||||||
Db    421  IVNEARIPACSL  432
```

RESULT 13
US-11-056-354-4
; Sequence 4, Application US/11056354
; Publication No. US20050246780A1
; GENERAL INFORMATION:
; APPLICANT: SHORT, Jay M.
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin A.
; APPLICANT: BARTON, Nelson R.
; APPLICANT: GARRETT, James B.
; APPLICANT: O'DONOGHUE, Eileen
; APPLICANT: MATHUR, Eric J.
; TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM AND METHODS
; TITLE OF INVENTION: FOR MAKING AND USING THEM
; FILE REFERENCE: 564462001803/D1370-9C1
; CURRENT APPLICATION NUMBER: US/11/056,354
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 10/156,660
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 4
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli appA phytase
; SOFTWARE: Patentin version 3.1

US-11-056-354-4

Query Match 97.1%; Score 2235; DB 6; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.7e-187;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60

Qy     61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120

Qy    121  APAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  APAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180

Qy    181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240

Qy    241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300

Qy    301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360

Qy    241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300

Qy    361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
```

```
Qy   421  IVNEARIPACSL 432
          ||||||||||||
Db   421  IVNEARIPACSL 432
```

RESULT 14
US-11-018-709-1
Sequence 1, Application US/11018709
Publication No. US20050095691A1
GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
; FILE REFERENCE: 19603/4031
; CURRENT APPLICATION NUMBER: US/11/018,709
; CURRENT FILING DATE: 2004-12-20
; PRIOR APPLICATION NUMBER: US/09/715,477
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli

US-11-018-709-1

Query Match   96.9%; Score 2230; DB 6; Length 432;
Best Local Similarity   99.3%; Pred. No. 4.6e-187;
Matches 429; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

```
Qy     1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60

Qy    61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy   121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy   181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy   241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy   301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy   361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy   421  IVNEARIPACSL 432
          ||||||||||||
Db   421  IVNEARIPACSL 432
```

RESULT 15
US-10-284-962-3
Sequence 3, Application US/10284962
Publication No. US20030206913A1
GENERAL INFORMATION:
; APPLICANT: Webel, Douglas M.
; APPLICANT: Orr, Donald E.
; APPLICANT: Ruch, Frank E.
; TITLE OF INVENTION: PHYTASE-CONTAINING ANIMAL FOOD AND METHOD
; FILE REFERENCE: 834460-717225
; CURRENT APPLICATION NUMBER: US/10/284,962
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/335,303
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli

US-10-284-962-3

Query Match   96.5%; Score 2221; DB 4; Length 432;
Best Local Similarity   99.1%; Pred. No. 2.8e-186;
Matches 428; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

```
Qy     1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60

Qy    61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQPGQVAIIADVDERTRKTGE 120

Qy   121  AFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy   181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy   241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy   301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy   361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy   421  IVNEARIPACSL 432
          ||||||||||||
Db   421  IVNEARIPACSL 432
```

Search completed: June 13, 2006, 10:50:37
Job time : 145.303 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2006, 10:28:49 ; Search time 36.3303 Seconds
                                       (without alignments)
                                       1060.093 Million cell updates/sec

Title: US-10-601-319-2
Perfect score: 2302
Sequence: 1 MKAILIPFLSLLIPLTPQSA.........IVNEARIPACSLRSHHHHHH 440
Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5
Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued Patents AA:*
                 1: /EMC_Celerra_SIDS3/ptodata/2/iaa/5_COMB.pep:*
                 2: /EMC_Celerra_SIDS3/ptodata/2/iaa/6_COMB.pep:*
                 3: /EMC_Celerra_SIDS3/ptodata/2/iaa/7_COMB.pep:*
                 4: /EMC_Celerra_SIDS3/ptodata/2/iaa/H_COMB.pep:*
                 5: /EMC_Celerra_SIDS3/ptodata/2/iaa/PCTUS_COMB.pep:*
                 6: /EMC_Celerra_SIDS3/ptodata/2/iaa/RB_COMB.pep:*
                 7: /EMC_Celerra_SIDS3/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2302	100.0	440	2	US-09-259-214-2	Sequence 2, Appli
2	2302	100.0	440	2	US-10-083-452-8	Sequence 8, Appli
3	2302	100.0	440	2	US-09-291-931-2	Sequence 2, Appli
4	2302	100.0	440	2	US-09-580-515-2	Sequence 2, Appli
5	2235	97.1	432	2	US-09-866-379-2	Sequence 2, Appli
6	2235	100.0	440	2	US-09-866-379-8	Sequence 8, Appli
7	2230	96.9	432	2	US-09-715-477-1	Sequence 1, Appli
8	2221	96.5	432	2	US-10-266-041A-1	Sequence 1, Appli
9	2221	96.5	433	2	US-09-540-149A-1	Sequence 1, Appli
10	2210	96.0	432	2	US-09-715-477-3	Sequence 3, Appli
11	2210	96.0	423	1	US-08-910-798-10	Sequence 10, Appl
12	2141	93.0	430	2	US-09-866-379-10	Sequence 10, Appl
13	528	22.9	421	2	US-09-489-039A-7512	Sequence 7512, Ap
14	376	16.3	522	2	US-09-489-039A-13501	Sequence 13501, A
15	119	5.2	439	2	US-09-044-718-3	Sequence 3, Appli
16	119	5.2	439	2	US-10-062-848-3	Sequence 3, Appli
17	119	5.2	440	2	US-09-684-855-106	Sequence 106, App
18	119	5.2	440	2	US-09-684-855-128	Sequence 128, App
19	119	5.2	440	2	US-09-488-265B-6	Sequence 6, Appli
20	119	5.2	440	2	US-09-488-265B-6	Sequence 151, App
21	119	5.2	449	2	US-09-044-718-12	Sequence 12, Appl
22	119	5.2	449	2	US-10-062-848-12	Sequence 12, Appl
23	119	5.2	465	2	US-08-868-435-33	Sequence 33, Appl
24	119	5.2	465	2	US-09-744-231-33	Sequence 33, Appl
25	119	5.2	465	2	US-09-744-718-78	Sequence 78, Appl
26	119	5.2	465	2	US-09-636-499-6	Sequence 6, Appli
27	119	5.2	465	2	US-09-273-871A-8	Sequence 8, Appli
28	119	5.2	465	2	US-10-083-452-8	Sequence 8, Appli
29	119	5.2	465	2	US-09-635-504-33	Sequence 33, Appl
30	119	5.2	465	3	US-10-062-848-78	Sequence 78, Appl
31	119	5.2	465	3	US-10-229-358-6	Sequence 6, Appli
32	118	5.1	440	2	US-09-684-855-107	Sequence 107, App
33	118	5.1	440	2	US-09-684-855-129	Sequence 129, App
34	118	5.1	440	2	US-09-684-855-152	Sequence 152, App
35	118	5.1	440	2	US-09-488-265B-7	Sequence 7, Appli
36	117	5.1	386	1	US-08-758-213-1	Sequence 1, Appli
37	117	5.1	386	1	US-08-692-787-48	Sequence 48, Appl
38	117	5.1	386	1	US-09-097-199-48	Sequence 48, Appl
39	117	5.1	386	1	US-09-949-016-6022	Sequence 6022, Ap
40	117	5.1	440	2	US-08-579-823A-2	Sequence 2, Appli
41	117	5.1	515	2	US-09-146-283-2	Sequence 2, Appli
42	117	5.1	515	2	US-09-344-195-2	Sequence 2, Appli
43	117	5.1	515	2	US-09-270-767-41623	Sequence 41623, A
44	114.5	5.0	326	3	US-10-144-678A-1020	Sequence 1020, Ap
45	111	4.8	440	2	US-09-684-855-108	Sequence 108, App

ALIGNMENTS

RESULT 1
US-09-259-214-2
Sequence 2, Application US/09259214A
; Patent No. 6110719
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: DIVER1370-1
; CURRENT APPLICATION NUMBER: US/09/259,214A
; CURRENT FILING DATE: 1999-03-01
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-259-214-2

Query Match   100.0%; Score 2302; DB 2; Length 440;
Best Local Similarity 100.0%; Pred. No. 3.9e-234;
Matches 440; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTKTGE 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTKTGE 120

Qy   121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy   241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVERW 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVERW 360

Qy   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

```
Db   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   421 IVNEARIPACSLRSHHHHHH 440
         ||||||||||||||||||||
Db   421 IVNEARIPACSLRSHHHHHH 440
```

RESULT 2
US-09-318-528-2
; Sequence 2, Application US/09318528
; Patent No. 6183740
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/318,528
; CURRENT FILING DATE: 1999-05-25
; EARLIER APPLICATION NUMBER: 09/291,931
; EARLIER FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-318-528-2

  Query Match   100.0%; Score 2302; DB 2; Length 440;
  Best Local Similarity   100.0%; Pred. No. 3.9e-234;
  Matches 440; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy     1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESSVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESSVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy    61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy   121 APFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 APFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy   241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW 360

Qy   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy   421 IVNEARIPACSLRSHHHHHH 440
         ||||||||||||||||||||
Db   421 IVNEARIPACSLRSHHHHHH 440
```

RESULT 3
US-09-291-931-2
; Sequence 2, Application US/09291931A
; Patent No. 6190897

GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/291,931A
; CURRENT FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-291-931-2

  Query Match   100.0%; Score 2302; DB 2; Length 440;
  Best Local Similarity   100.0%; Pred. No. 3.9e-234;
  Matches 440; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy     1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESSVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESSVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy    61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy   121 APFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 APFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy   241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW 360

Qy   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy   421 IVNEARIPACSLRSHHHHHH 440
         ||||||||||||||||||||
Db   421 IVNEARIPACSLRSHHHHHH 440
```

RESULT 4
US-09-580-515-2
; Sequence 2, Application US/09580515
; Patent No. 6720014
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/580,515
; CURRENT FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01

; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
;
; SEQ ID NO 2
;   LENGTH: 440
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-09-580-515-2

Query Match   100.0%; Score 2302; DB 2; Length 440;
Best Local Similarity   100.0%; Pred. No. 3.9e-234;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy   121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy   241 EIFLLQQAQGMPEPGWGKRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 EIFLLQQAQGMPEPGWGKRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360

Qy   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy   421 IVNEARIPACSLRSHHHHHH 440
          ||||||||||||||||||||
Db   421 IVNEARIPACSLRSHHHHHH 440

RESULT 5
US-09-866-379-2
; Sequence 2, Application US/09866379
; Patent No. 6855365
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; FILE REFERENCE: DIVER1370-7
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; APPLICANT: O'DONOGHUE, Eileen
; APPLICANT: GARRETT, James
; APPLICANT: BARTON, Nelson
; APPLICANT: GRAY, Kevin
; APPLICANT: KRETZ, Keith
; APPLICANT: SHORT, Jay
; APPLICANT: DIVERSA CORPORATION
; GENERAL INFORMATION:
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: US 08/910,798
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1

; SEQ ID NO 2
;   LENGTH: 440
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-09-866-379-2

Query Match   100.0%; Score 2302; DB 2; Length 440;
Best Local Similarity   100.0%; Pred. No. 3.9e-234;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy   121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy   241 EIFLLQQAQGMPEPGWGKRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 EIFLLQQAQGMPEPGWGKRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360

Qy   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy   421 IVNEARIPACSLRSHHHHHH 440
          ||||||||||||||||||||
Db   421 IVNEARIPACSLRSHHHHHH 440

RESULT 6
US-09-866-379-8
; Sequence 8, Application US/09866379
; Patent No. 6855365
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; FILE REFERENCE: DIVER1370-7
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; APPLICANT: O'DONOGHUE, Eileen
; APPLICANT: GARRETT, James
; APPLICANT: BARTON, Nelson
; APPLICANT: GRAY, Kevin
; APPLICANT: KRETZ, Keith
; APPLICANT: SHORT, Jay
; APPLICANT: DIVERSA CORPORATION
; GENERAL INFORMATION:
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: US 08/910,798
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 8
;   LENGTH: 432

```
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-866-379-8

Query Match        97.1%;  Score  2235;  DB 2;  Length 432;
Best Local Similarity   99.5%;  Pred. No. 4.5e-227;
Matches  430;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60

Qy   61  TWPVKLGWLTPRGGELIAYLGHYQRORLVADGLLAKKKGCPQSGQVAIIADVDERTRKTGE  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  TWPVKLGWLTPRGGELIAYLGHYQRORLVADGLLAKKKGCPQSGQVAIIADVDERTRKTGE  120

Qy  121  AFAAGLAPDCAITVHTQADTSSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  AFAAGLAPDCAITVHTQADTSSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180

Qy  181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240

Qy  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA  300

Qy  301  LTPHPPQOKQAYGVTLPFTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  LTPHPPQOKQAYGVTLPFTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420

Qy  421  IVNEARIPACSL  432
         ||||||||||||
Db  421  IVNEARIPACSL  432


RESULT 7
US-09-715-477-1
; Sequence 1, Application US/09715477
; Patent No. 6841370
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
; FILE REFERENCE: 19603/4031
; CURRENT APPLICATION NUMBER: US/09/715,477
; CURRENT FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-715-477-1

Query Match        96.9%;  Score  2230;  DB 2;  Length 432;
Best Local Similarity   99.3%;  Pred. No. 1.5e-226;
Matches  429;  Conservative  0;  Mismatches  3;  Indels  0;  Gaps  0;

Qy    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60

Qy   61  TWPVKLGWLTPRGGELIAYLGHYQRORLVADGLLAKKKGCPQSGQVAIIADVDERTRKTGE  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  TWPVKLGWLTPRGGELIAYLGHYQRORLVADGLLAKKKGCPQSGQVAIIADVDERTRKTGE  120

Qy  121  AFAAGLAPDCAITVHTQADTSSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  AFAAGLAPDCAITVHTQADTSSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180

Qy  181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240

Qy  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA  300

Qy  301  LTPHPPQOKQAYGVTLPFTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  LTPHPPQOKQAYGVTLPFTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420

Qy  421  IVNEARIPACSL  432
         ||||||||||||
Db  421  IVNEARIPACSL  432
```

```
Qy  121  AFAAGLAPDCAITVHTQADTSSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  AFAAGLAPDCAITVHTQADTSSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180

Qy  181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240

Qy  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA  300

Qy  301  LTPHPPQOKQAYGVTLPFTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  LTPHPPQOKQAYGVTLPFTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420

Qy  421  IVNEARIPACSL  432
         ||||||||||||
Db  421  IVNEARIPACSL  432


RESULT 8
US-10-266-041A-1
; Sequence 1, Application US/10266041A
; Patent No. 6974690
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: PHOSPHATASES WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2798
; CURRENT APPLICATION NUMBER: US/10/266,041A
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: 60/127,032
; PRIOR FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 09/540,149
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-266-041A-1

Query Match        96.5%;  Score  2221;  DB 2;  Length 432;
Best Local Similarity   99.1%;  Pred. No. 1.4e-225;
Matches  428;  Conservative  0;  Mismatches  4;  Indels  0;  Gaps  0;

Qy    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60

Qy   61  TWPVKLGWLTPRGGELIAYLGHYQRORLVADGLLAKKKGCPQSGQVAIIADVDERTRKTGE  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  TWPVKLGWLTPRGGELIAYLGHYQRORLVADGLLAKKKGCPQSGQVAIIADVDERTRKTGE  120

Qy  121  AFAAGLAPDCAITVHTQADTSSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  AFAAGLAPDCAITVHTQADTSSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180

Qy  181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
         |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db  181  RQTAFRELERVLNFSQLNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240

Qy  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300

Qy  301  LTPHPPQOKQAYGVTLPFTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  LTPHPPQOKQAYGVTLPFTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
```

RESULT 9
US-09-540-149A-1

; Sequence 1, Application US/09540149A
; Patent No. 6511699
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2791
; CURRENT APPLICATION NUMBER: US/09/540,149A
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 60/127,032
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 1
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (433)
; OTHER INFORMATION: Xaa at position 433 in this sequence is unknown
US-09-540-149A-1

Query Match   96.5%; Score 2221; DB 2; Length 433;
Best Local Similarity 99.1%; Pred. No. 1.4e-225;
Matches 428; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy      1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy     61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy    121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy    181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy    241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy    301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360

Qy    361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy    421 IVNEARIPACSL 432
            ||||||||||||
Db    421 IVNEARIPACSL 432
```

RESULT 10

US-09-715-477-3

; Sequence 3, Application US/09715477
; Patent No. 6841370
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
; FILE REFERENCE: 19603/4031
; CURRENT APPLICATION NUMBER: US/09/715,477
; CURRENT FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-715-477-3

Query Match   96.0%; Score 2210; DB 2; Length 432;
Best Local Similarity 98.6%; Pred. No. 2e-224;
Matches 426; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy      1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy     61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy    121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy    181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
            |||||||||||||||||||||:|||::|:||||||||||||||||||||||||||||||
Db    181 RQTAFRELERVLNFPQSNLNLKREKQNESCNLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy    241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy    301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360

Qy    361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy    421 IVNEARIPACSL 432
            ||||||||||||
Db    421 IVNEARIPACSL 432
```

RESULT 11

US-08-910-798-2

; Sequence 2, Application US/08910798
; Patent No. 5876997
; GENERAL INFORMATION:
; APPLICANT: KRETZ
; TITLE OF INVENTION: NOVEL PHYTASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
;

```
  ;    COMPUTER: IBM PC compatible
  ;    OPERATING SYSTEM: PC-DOS/MS-DOS
  ;    SOFTWARE: Patentin Release #1.0, Version #1.25
  ;    CURRENT APPLICATION DATA:
  ;        APPLICATION NUMBER: US/08/910,798
  ;        FILING DATE: August 13,1997
  ;        CLASSIFICATION: 435
  ;    ATTORNEY/AGENT INFORMATION:
  ;        NAME: HAILE, PH.D., LISA A.
  ;        REGISTRATION NUMBER: 38,347
  ;        REFERENCE/DOCKET NUMBER: 09010/029001
  ;    TELECOMMUNICATION INFORMATION:
  ;        TELEPHONE: 619/678-5070
  ;        TELEFAX: 619/678-5099
  ;    INFORMATION FOR SEQ ID NO: 2:
  ;    SEQUENCE CHARACTERISTICS:
  ;        LENGTH: 423 amino acids
  ;        TYPE: amino acid
  ;        TOPOLOGY: linear
  ;    MOLECULE TYPE: protein
  US-08-910-798-2
```

Query Match   95.3%;   Score 2194.5;   DB 1;   Length 423;
Best Local Similarity   96.1%;   Pred. No. 8.3e-223;
Matches   423;   Conservative   0;   Mismatches   0;   Indels   17;   Gaps   1;

```
Qy    1 MKAILIPFLSLLLIPLITPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLLIPLITPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTG- 119

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  120 -----------QADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 163

Qy  181 RQTAFRELERVLNFPQSNLCLKREKQODESCCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  164 RQTAFRELERVLNFPQSNLCLKREKQODESCCSLTQALPSELKVSADNVSLTGAVSLASMLT 223

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  224 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 283

Qy  301 LTPHPPQKQAYGVTLPTSVLPIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  284 LTPHPPQKQAYGVTLPTSVLPIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 343

Qy  361 RRLSDNSQWIQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAOGMCSLAGFTQ 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  344 RRLSDNSQWIQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAOGMCSLAGFTQ 403

Qy  421 IVNEARIPACSLRSHHHHHH 440
        ||||||||||||||
Db  404 IVNEARIPACSLRSHHHHHH 423
```

RESULT 12
US-09-866-379-10
Sequence 10, Application US/0866379
Patent No. 6855365

```
  ;    GENERAL INFORMATION:
  ;    APPLICANT: DIVERSA CORPORATION
  ;    APPLICANT: SHORT, Jay
  ;    APPLICANT: KRETZ, Keith
  ;    APPLICANT: GRAY, Kevin
  ;    APPLICANT: BARTON, Nelson
  ;    APPLICANT: GARRETT, James
  ;    APPLICANT: O'DONOGHUE, Eileen
  ;    TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
  ;    FILE REFERENCE: DIVER1370-7
```

```
  ;    CURRENT APPLICATION NUMBER: US/09/866,379
  ;    CURRENT FILING DATE: 2001-05-24
  ;    PRIOR APPLICATION NUMBER: US 09/866,379
  ;    PRIOR FILING DATE: 2000-05-25
  ;    PRIOR APPLICATION NUMBER: US 09/580,515
  ;    PRIOR FILING DATE: 1999-05-25
  ;    PRIOR APPLICATION NUMBER: US 09/318,528
  ;    PRIOR FILING DATE: 1999-04-13
  ;    PRIOR APPLICATION NUMBER: US 09/291,931
  ;    PRIOR FILING DATE: 1999-03-01
  ;    PRIOR APPLICATION NUMBER: US 09/259,214
  ;    PRIOR FILING DATE: 1997-08-13
  ;    PRIOR APPLICATION NUMBER: US 08/910,798
  ;    NUMBER OF SEQ ID NOS: 10
  ;    SOFTWARE: Patentin version 3.1
  ;    SEQ ID NO 10
  ;        LENGTH: 430
  ;        TYPE: PRT
  ;        ORGANISM: Artificial Sequence
  ;        FEATURE:
  ;        OTHER INFORMATION: Modified phytase
  US-09-866-379-10
```

Query Match   93.0%;   Score 2141;   DB 2;   Length 430;
Best Local Similarity   97.2%;   Pred. No. 3.9e-217;
Matches   420;   Conservative   0;   Mismatches   10;   Indels   2;   Gaps   2;

```
Qy    1 MKAILIPFLSLLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAIL-EAGGSIADFTGH 179

Qy  181 RQTAFRELERVLNFPQSNLCLKREKQODESCCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  180 YQTAFRELERVLNFPQSNLCLKREKQODESCCSLTQALPSELKVSADCVSLTGAVSLASMLT 239

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  240 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFD-LQRTPEVARSRATPLLDLIKTA 298

Qy  301 LTPHPPQKQAYGVTLPTSVLPIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  299 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 358

Qy  361 RRLSDNSQWIQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAOGMCSLAGFTQ 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  359 RRLSDNSQWIQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAOGMCSLAGFTQ 418

Qy  421 IVNEARIPACSL 432
        ||||||||||||
Db  419 IVNEARIPACSL 430
```

RESULT 13
US-09-489-039A-7512
Sequence 7512, Application US/0489039A
Patent No. 6610836

```
  ;    GENERAL INFORMATION:
  ;    APPLICANT: Gary Breton et. al
  ;    TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
  ;    TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
  ;    FILE REFERENCE: 2709.2004001
  ;    CURRENT APPLICATION NUMBER: US/09/489,039A
  ;    CURRENT FILING DATE: 2000-01-27
  ;    PRIOR APPLICATION NUMBER: US 60/117,747
  ;    PRIOR FILING DATE: 1999-01-29
  ;    NUMBER OF SEQ ID NOS: 14342
```

```
; SEQ ID NO 7512
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae

Best Match
Query Match       22.9%;  Score 528;  DB 2;  Length 421;
Matches 134;  Conservative 69;  Mismatches 184;  Indels 30;  Gaps 10;
US-09-489-039A-7512


Qy   19  SAFAQSE------PE-LKLESVVIVSRHGVRAP-TKATQLMQDVTPDAWPTWPVKLGWLTPR   72
Db   23  SAGAQAQDKAAPEGYQLQQVLIMSRHNLRAPLANNNGSVLBQSTAKANPQWDVPGGQLTTK   82

Qy   73  GGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGEAFAAGLAPDCAI  132
Db   83  GGVLEVVMGHYMREWLAQQKLVTSGBCPPENAVTAYANSLQRTVATAQFFITGAFPGCGI  142

Qy  133  TVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAI-LSRAGGSIADFPGHRQTAFRELERV  191
Db  143  PVHHQPQMGTMDPTFNPVITDDSPAFREKALQGAMEKRQGMQLTE------SYKLLETM  195

Qy  192  LNFPQOSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLTEIFLLQQAQGM  251
Db  196  IDYRNSPSC----KEKKVCSLSEGKDTPSAGYQQEPGVSGPLKVGNSLVDAFTLQYYEGF  251

Qy  252  P--EPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAALTPHPPQKQ  309
Db  252  PKDQVAWGBIASDKQWRVLSKLKNGYQDSLFTSVAVAQNVAKPLVKYIDNALVGBGAGSK-  310

Qy  310  AYGVTLPTSVLFIAGHDTNLANLGGALELN-WTLPGQPDNTPPGGELVFERWRRLSDNSQ  368
Db  311  ------AKVLLLVGHSNIASLLTALDFKPYQLPGGQYERTPIGGKLLFQRWHDSAGNRD  363

Qy  369  WIQVSLVFQTLQQMRDKTPLSLNTPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEA  425
Db  364  LMKIEVYYQSTEQLRNADALTLQAPFQRVTLALNGCPV-DDQGPCPLETFKMVINEA  419


RESULT 14
US-09-489-039A-13501
Sequence 13501, Application US/0948039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13501
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13501

Query Match       16.3%;  Score 376;  DB 2;  Length 522;
Best Local Similarity 25.8%;  Pred. No. 1.8e-30;
Matches 115;  Conservative 75;  Mismatches 209;  Indels 46;  Gaps 15;


Qy    2  KAILIPFLSLILPLTP-QSAFAQSEPELKLESVVIVSRHGVRAP-TKAT-QLMQDVTPDAW   59
Db  107  QGLLRLFIACALPLLALQSAAA----ADWQLEKVVELSRHGIRPPTAGNREAIBAAGRPW  163

Qy   60  PTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTG  119
Db  164  TEWTTHDGELTGHGYAAVVNNGRAEGQHYRQLGLL-QAGCPTAESIVRASPLQRTRATA  222

Qy  120  EAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTG  179
Db  223  QALVDGAFPGCGVAIHYV--SGDADPLFQTDKFAATQTDPARQLAAVKEKAG----DLAQ  276


RESULT 15
US-09-044-718-3
Sequence 3, Application US/09044718
; Patent No. 6339605
; GENERAL INFORMATION:
; APPLICANT: KOSTREWA, Dirk
; APPLICANT: PASAMONTES, Luis
; APPLICANT: TOMSCHY, Andrea
; APPLICANT: van LOON, Adolphus
; APPLICANT: VOGEL, Kurt
; APPLICANT: WYSS, Markus
; TITLE OF INVENTION: MODIFIED PHYTASES
; FILE REFERENCE: Modified Phytases
; CURRENT APPLICATION NUMBER: US/09/044,718
; CURRENT FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: EP 97810175.6
; PRIOR FILING DATE: 1997-03-25
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-09-044-718-3

Query Match        5.2%;  Score 119;  DB 2;  Length 439;
Best Local Similarity 21.7%;  Pred. No. 0.002;
Matches 102;  Conservative 70;  Mismatches 183;  Indels 114;  Gaps 24;


Qy    7  PFLSLILPLTPQSAFAQSEPELKLESVVIVSRHGVRAPT-----KATQLMQDVTPDA--   58
Db   26  PFFSLEDELSVSSKLPK---DCRITLVQVLSRHGARYPTSSKSKKYKLVTAIQANATDF   82

Qy   59  ------WPTWPVKLGW--LTPRGGELIAYLG--HYQRQRLVADGLLAKKGCPQSGQVAII  108
Db   83  KGKFAFLKTYNYTLGADDLTPFGEQQLVNSGIKFYQRYKALARSVVP--------FIR  132

Qy  109  ADVDERTRKTGEAFAAGL-----------APDCAITVHTQADTSSPDPLFNPLKTGV  154
Db  133  ASGSDRVIASGEKFIEGFQQQAKLADPGATNRAAPAISVII-PESETN-----NTLDHGV  186

Qy  155  CQLDNANVTDAILSRAGGSSI-ADFTGHRQTAFRELERVLNFPQOSNLCLKREKQDESCSLT  213
Db  187  C-------TKFEASQLGDEVAANFT----------------DIRARAEKHLPGVTLT  225

Qy  214  QALPSEL--KVSADNVSLTGAVSLASMLTEIFLLQQAGMEPEPGWGRITDSHQWNTLLSL  271
Db  226  DEDVVSLMDMCSFDTVARTSDASQLSPFCQLFTHNE-------W-----KKYNYLQSL  271

Qy  272  HNAQFYLLQRTPEVARSRATPLLDLIMAALTPHPPQKQ-------------AYGVTLPTSVLFIA  323
```

```
Db  272  --GKYYGYGAGNPLGPAQGIGPTNELIARLTRSPVQDHTSTNSTLVSNPATPPLNATMYV 329
             ::|            :     :: | ::|  ||                  || :
Qy  324  --GHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERWRRLSD-NSQWI---QVSLVFQ 377
           |  : :: ||          |  :    |:     ||  |:          |:
Db  330  DFSHDNSMVSIFPAL-------GLYNGTBPLSRTSVBSAKELDGYSASWVVPFGARAYFE 382

Qy  378  TLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEAR         426
         |:|  ||| | :  |:     :  |  : | ::::||
Db  383  TMQCKSEKEPL-VRALINDRVVPLHGCDV-DKLGRCKLNDFVKGLSWAR         429
```

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

Run on: June 14, 2006, 11:48:02 ; Search time 4986.33 Seconds
                                             (without alignments)
                                             7401.587 Million cell updates/sec

Title: US-10-601-319-2
Perfect score: 2302
Sequence: 1 MKAILIPFLSLLIPLTPQSA.........IVNEARIPACSLRSHHHHHH 440

Scoring table: BLOSUM62
                  Xgapop 10.0 , Xgapext   0.5
                  Ygapop 10.0 , Ygapext   0.5
                  Fgapop   6.0 , Fgapext   7.0
                  Delop   6.0 , Delext   7.0

Searched: 48236798 seqs, 2795965780 residues

Total number of hits satisfying chosen parameters: 96473596

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/abs//ABSWEB_spool/US10601319//runat_12062006_180052_27516/app_query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPUT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=LOCAL
-USER=US10601319 @CGN_1.1_12067 @runat_12062006_180052_27516 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -YGAPOP=10 -YGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7 -FGAPOP=6 -FGAPEXT=7

Database :   EST:*
                  1: gb_est1:*
                  2: gb_est3:*
                  3: gb_est4:*
                  4: gb_est5:*
                  5: gb_est6:*
                  6: gb_htc:*
                  7: gb_est2:*
                  8: gb_est7:*
                  9: gb_est8:*
                 10: gb_est9:*
                 11: gb_gss1:*
                 12: gb_gss2:*
                 13: gb_gss3:*
                 14: gb_gss4:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	1044	45.4	616	2	BJ074127	BJ074127 BJ074127
c 2	1036	45.0	853	13	CL662734	CL662734 PRI0142b
c 3	913	39.7	529	3	BU713770	BU713770 SJAABUG02
c 4	896	38.9	746	4	CA093060	CA093060 SCCCL200

c 5	846	36.8	595	5	CP326092	CP326092 JMT1--05-
c 6	838	36.4	895	8	CN762997	CN762997 ID0AAA5DC
c 7	767	33.3	868	8	CN754382	CN754382 ID0AAA12D
c 8	757	32.9	864	6	CN757565	CN757565 ID0AAA1DF
c 9	643	27.9	383	7	AW036132	AW036132 EST274508
c 10	643	27.9	383	7	AW036134	AW036134 EST274510
c 11	641	27.8	393	7	BE520240	BE520240 M1B12STM
c 12	591.5	25.7	531	3	BU713769	BU713769 SJAABUG01
c 13	522	22.7	706	8	CN759004	CN759004 ID0AAA24B
c 14	511.5	22.2	354	1	AA545747	AA545747 HBMSF1B4-
c 15	442	19.2	707	2	BJ619443	BJ619443 BJ619443
c 16	392	17.0	320	7	BE436403	BE436403 EST407481
c 17	374.5	16.3	1175	2	CX943192	CX943192 LamjGestG
c 18	355	15.4	208	2	BM409016	BM409016 EST583343
c 19	355	15.4	208	2	BM412806	BM412806 EST587133
c 20	350.5	15.2	821	13	CL665831	CL665831 PRI014d_E
c 21	345.5	15.0	842	13	CL665745	CL665745 PRI014c_E
c 22	339	14.7	546	2	BJ040796	BJ040796 BJ040796
c 23	291	12.6	757	4	CA093061	CA093061 SCCCL200
c 24	287	12.5	657	2	BG457132	BG457132 NF100COSP
c 25	284	12.3	1171	2	CV672454	CV672454 RET7SJ_01
c 26	278	12.1	179	7	BB461872	BB461872 EST413291
c 27	269	11.7	508	2	BJ030203	BJ030203 BJ030203
c 28	228.5	9.9	480	1	AL586973	AL586973 AL586973
c 29	225	9.8	254	7	BE520241	BE520241 M1B12XTM
c 30	225	9.8	470	9	DN201839	DN201839 USDA-FP_1
c 31	219.5	9.5	692	6	CK564106	CK564106 rswpb0_00
c 32	215	9.3	772	14	DX010757	DX010757 KBrB00lC0
c 33	195	8.5	551	6	CK535965	CK535965 rswgb0_00
c 34	186	8.1	692	9	CX700688	CX700688 Hm_L6w_04
c 35	185.5	8.1	575	5	CK543282	CK543282 rswhb0_01
c 36	183	7.9	687	10	DV079873	DV079873 I13SFmidg
c 37	172.5	7.5	696	1	AU004475	AU004475 AU004475
c 38	172	7.5	734	5	CK507686	CK507686 rBwda0_00
c 39	163	7.1	769	8	CO386885	CO386885 AGENCOURT
c 40	153	6.5	863	9	CV672453	CV672453 RET7SJ_01
c 41	150	6.5	1044	6	BX021319	BX021319 Single_re
c 42	144.5	6.3	645	2	CN508OM3	CN508OM3 CN508OM3
c 43	139.5	6.1	1430	2	BJ037937	BJ037937 full-leng
c 44	139.5	6.0	2201	6	CR615614	CR615614 Pongo pyg
c 45	137.5	6.0	2201	6	CR926080	CR926080 Pongo pyg
	133	5.8	1125	10	DW035689	DW035689 CFW261-GO

ALIGNMENTS

RESULT 1
BJ074127/c
LOCUS   BJ074127   616 bp   mRNA   linear
DEFINITION   BJ074127 NIBB Mochii normalized Xenopus tailbud library Xenopus
               laevis cDNA clone XL090j06 5', mRNA sequence.
ACCESSION   BJ074127
VERSION   BJ074127.1 GI:17504316
KEYWORDS   EST.
SOURCE   Xenopus laevis (African clawed frog)
  ORGANISM   Xenopus laevis
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
               Xenopodinae; Xenopus; Xenopus.
REFERENCE   1 (bases 1 to 616)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and
               Kohara,Y.
  TITLE   Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT   Contact: Tadasu Shin-i
               Center For Genetic Resource Information
               National Institute of Genetics
               1111 Yata, Mishima, Shizuoka 411-8540, Japan
               Tel: 81-559-81-6856
               Fax: 81-559-81-6855
               Email: tshini@genes.nig.ac.jp
               The information of this clone is available through the following
               URL.
               http://xenopus.nibb.ac.jp.

FEATURES   Location/Qualifiers
  source   1. .616
                /organism="Xenopus laevis"
                /mol_type="mRNA"
                /clone="XLO90J06"
                /db_xref="taxon:8355"
                /tissue_type="whole embryo"
                /dev_stage="stage 25"
                /clone_lib="NIBB Mochii normalized Xenopus tailbud library"

ORIGIN

US-10-601-319-2 (1-440) x BJ074127 (1-616)

Alignment Scores:
Pred. No.:
Score: 5.31e-96
                          1044.00   Length: 616
Percent Similarity: 99.5%   Matches: 204
Best Local Similarity: 99.5%   Conservative: 0
Query Match: 2   Mismatches: 1
                                               Indels: 0
DB: Gaps: 0

Qy   89 ValAlaAspGlyLeuLeuAlaLeuLysGlyGlyCysProGlnSerGlyGlnValAlaAlaIleIle 108
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 615 GTAGCCGACGGATTGCTGGGCAGAAAAAGGGGTCGCCCCAGTCTGGGTCAGGTCGCGGATTATT 556

Qy 109 AlaAspValAspGluArgThrArgLysThrGlyIleValPheAlaAlaGlyIleuAlaPro 128
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 555 GCTGATGTCGACGAGCGTACCCGTAAAAACAGGCGAAGCCTTCGCCCGGGGCTGGCACCT 496

Qy 129 AspCysAlaIleThrValHisThrGlnAlaAspThrSerSerProAspProLeuPheAsn 148
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 495 GACTGTGCAATAACCCTACATACCCAGGCAGATACGTCCCAGTCCCGATCCGTTATTTAAT 436

Qy 149 ProLeuLysThrGlyValCysGlnLeuuAspAsnAlaAsnValThrAspAlaIleIleLeuSer 168
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 435 CCTCTAAAAAACTGGCGTTTGCCAACTGGGATAACGCGATCAACTGAACGCGATCCTCAGC 376

Qy 169 ArgAlaGlyGlySerIleAlaAspPheThrGlyHisArgGlnThrAlaPheArgGlyGluLeu 188
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 375 AGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCATCGGCAAACGGCCGTTTCGGCGAACTG 316

Qy 189 GluArgValLeuAsnPheProGlnSerArgAsnLeuCysLeuLysArgGlulysGlnAspGlu 208
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 315 GAACCGGTGCTTAATTTTCCCGCAATCAAAACTTGTGCCTTAAACGTGAGAAACAGGACGAA 256

Qy 209 SerCysSerLeuThrGlnAlaLeuProSerGluLeuLysValSerAlaLeuAspAsnValSer 228
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 255 AGCTGTTCATTAAACCGAGGCAATTACCATCCGAACTCAAGGTGAGCGCCGACAATGTCTCA 196

Qy 229 LeuThrGlyAlaValSerLeuValAlaSerMetLeuuThrGlyIleuPheLeuLeuGlnGlnAla 248
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 195 TTAACCGGGTGCGGTAAGCCTCGCATCAATGCTGACCGGAGATATTTCTCCGGACAAGCA 136

Qy 249 GlnGlyMetProGlnProGlyTrpGlyArgIleThrAspSerHisGlnThrProAsnThrLeu 268
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 135 CTGGGAATGCCCGAGCCCGGGGTGGGGAAGGATCACCCAGTGAGAACACCTTG 76

Qy 269 LeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGluGlnArgThrProGluValAlaArgSer 288
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75 CTAAGTTTGCATAACCGCGAATTTATTTGCTACACACGCACGCCAGGGTTGCCCGCCAGC 16

Qy 289 ArgAlaThrProLeu 293
        |||||||||||||||
Db   15 CGCGCCACCCCGTTA 1

RESULT 2
LOCUS   CL662734
DEFINITION   PR10142b_B11 - PR10142b_B21 (853) Mixed stage fosmid library of P.
             pacificus var. California Pristionchus pacificus genomic, genomic
             survey sequence.
ACCESSION   CL662734

CL662734   853 bp   DNA   linear   GSS   09-JUL-2004

VERSION   CL662734.1 GI:50150877
KEYWORDS   GSS.
SOURCE   Pristionchus pacificus
  ORGANISM   Pristionchus pacificus
             Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
             Neodiplogasteridae; Pristionchus.

REFERENCE   1 (bases 1 to 853)
  AUTHORS   Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
  TITLE   AppDB: an AcedB database for the nematode satellite organism
             Pristionchus pacificus
  JOURNAL   Nucleic Acids Res. 32 (1), D421-D422 (2004)
  PUBMED   14681447

COMMENT   Contact: Sommer RJ
             Evolutionary Biology
             Max-Planck-Institute for Developmental Biology
             Spemannstr. 37-39, Tuebingen D-72076, Germany
             Tel: 0049707160137I
             Fax: 0049707160149B
             Email: ralf.sommer@tuebingen.mpg.de
             This library was generated at Caltech, Pasadena, USA and end
             sequenced at Vancouver, Canada.
             Seq primer: T7
             Class: fosmid ends.

FEATURES   Location/Qualifiers
  source   1. .853
                /organism="Pristionchus pacificus"
                /mol_type="genomic DNA"
                /strain="California"
                /db_xref="taxon:54126"
                /clone_lib="Mixed stage fosmid library of P. pacificus
                var. California"
                /note="Vector: pEpifos-5 Fosmid vector"

ORIGIN

US-10-601-319-2 (1-440) x CL662734 (1-853)

Alignment Scores:
Pred. No.:
Score: 5.73e-95
                          1036.00   Length: 853
Percent Similarity: 100.0%   Matches: 196
Best Local Similarity: 100.0%   Conservative: 0
Query Match: 13   Mismatches: 0
                                              Indels: 0
DB: Gaps: 0

Qy 237 SerMetLeuThrGluIlePheLeuLeuGluGlnAlaGlnGlyMetProGlyTrp 256
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2 TCAATGCTGACCGGAGATATTTCTCCGGACAAGCACAGGAGGAATGCCCGGACTGGTGG 61

Qy 257 GlyArgIleThrAspSerHisGlnThrProAsnThrLeuLeuSerLeuHisAsnAlaGlnPhe 276
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62 GGAAGGATCACCCAGTTCACACACAGTGGAACACCTTGCTAAGTTTGCATAACCGCGAATTT 121

Qy 277 TyrLeuLeuGlnArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspGlu 296
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122 TATTTGCTACACGCACGCCCAGGGTTGCCCGCCAGCGCGCGCCACCCCGTTATTGGATTTG 181

Qy 297 IleMetAlaAlaLeuThrProHisGlyHisAspThrAlaAsnLeuAlaAsnLeuGlyAlaLeu 316
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182 ATCATGGCAGCGTTGACGCCCCATCCACGGCAAAAACAGGGCGTATGGGGTGACATTACC 241

Qy 317 ThrSerValLeuPheIleAlaGlyHisAspThrAlaAsnLeuAlaAsnLeuGlyAlaLeu 336
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 242 ACTTCAGTACTGTTTATTGCCGGACACGATACTACTAATCTGGCAAACAGGG 301

Qy 337 GluLeuAsnTrpThrLeuProGlyGlnProAspAspAsnThrProGlyGlyGlyLeuVal 356
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 302 GAGCTCAACTGACGCTTCCCGGTCAGCCGGATAACACGCCCGGTGGTGGTGGACTGGTG 361

Qy 357 PheGluArgTrpArgLeuSerAspAsnSerGlnTrpIleGluIleValSerLeuLeuPhe 376
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 362 TTTGAACCGCTGGCGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGGCTGGTCTTC 421

Qy 377 GlnThrGluGlnGlnMetArgArgAspLysThrProLeuSerLeuAsnThrProProGlyGlu 396

```
Db   422  CAGACTTTACACGAGATGCGTGATAAAACGCCCGTGTCATTAAATAACCGCCGCCGGAGAG
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   397  ValLysLeuThrLeuAlaGlyCysGluGluArgAspAlaGlnGlyMetCysSerLeuAla  416
Db   482  GTGAAACTGACCCTGGCAGGATGTGAAGAGCGAGACGCGAAATGGCCAGGGCATGTGTTCGTTGGCA  541
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   417  GlyPheThrGlnHisValAsnGluAlaArgIleProAlaCysSerLeu  432
Db   542  GGTTTTACGCAAATCGTGAATGAAGCACGCATTACCGCGCGTGCAGTTTG  589
```

RESULT 3

```
Db         BU713770
LOCUS      BU713770              529 bp    mRNA    linear   EST 23-OCT-2003
DEFINITION SJAABUG02 Adult SjC 7/94 Schistosoma japonicum cDNA similar to
           pdb|1DKP|A Chain A, Crystal Structure Of Phytate Complex Of
           Escherichia Coli Phytase At Ph 6.6. Phytate Is Bound With Its
           3-Phosphate In The Active Site. Hg2+ Cation Acts As An
           Intermolecular Bridge, mRNA sequence.
ACCESSION  BU713770
VERSION    BU713770.1  GI:28321126
KEYWORDS   EST.
SOURCE     Schistosoma japonicum
  ORGANISM Schistosoma japonicum
           Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
           Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE  1  (bases 1 to 529)
  AUTHORS  Hu,W., Yan,Q., Shen,D.K., Liu,P., Zhu,Z.D., Song,H.D., Xu,X.R.,
           Wang,Z.J., Rong,Y.P., Zeng,L.C., Wu,J., Zhang,X., Wang,J.J.,
           Xu,X.N., Wang,S.Y., Fu,G., Zhang,X.L., Wang,Z.Q., Brindley,P.J.,
           McManus,D.P., Xue,C.L., Feng,Z., Chen,Z. and Han,Z.G.
  TITLE    Evolutionary and biomedical implications of a Schistosoma japonicum
           complementary DNA resource
  JOURNAL  Nat. Genet. 35 (2), 139-147 (2003)
   PUBMED  12973349
COMMENT    Contact: Zeguang Han
           Chinese National Human Genome Center at Shanghai
           351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
           201203, P. R. China
           Tel: 86-21-5001919(ex.45)
           Fax: 86-21-50801922
           Email: hanzg@chgc.sh.cn.
FEATURES             Location/Qualifiers
     source          1..529
                     /organism="Schistosoma japonicum"
                     /mol_type="mRNA"
                     /strain="Chinese (Anhui) strain"
                     /db_xref="taxon:6182"
                     /sex="Male and female"
                     /tissue_type="Whole body"
                     /dev_stage="Adult worms"
                     /lab_host="Mouse and rabbit"
                     /clone_lib="Adult SjC 7/94"
                     /note="Vector: Lambda ZAP-II XR.; Site 1: EcoR I; Site_2:
                     XhoI I; Several hundred adult Schistosoma japonicum
                     (Anhui, P.R. China, strain), of mixed sex, were perfused
                     from the mesenteries of experimentally infected mice and
                     rabbits at the Queensland Institute of Medical Research,
                     Brisbane, Australia (QIMR), and stored for several months
                     in liquid nitrogen. Subsequently, mRNA was isolated at the
                     QIMR from lysates of these worms by oligo dT
                     chromatography, using a kit from Pharmacia. The mRNA was
                     then shipped to Clontech, Palo Alto, CA, USA, who
                     constructed a cDNA library. First strand synthesis was
                     primed with an oligo-dT-XhoI-primer and synthesized using
                     M-MLV reverse transcriptase. Second strand synthesis was
                     accomplished with RNAse H and T4 DNA polymerase. The
                     double stranded cDNA was ligated to EcoRI linkers,
                     digested with EcoRI and XhoI, and ligated into the
                     phagemid vector lambda ZAP II XR. After construction of
                     this directional library by Clontech, it was returned to
                     the QIMR. During analysis of the library at the QIMR, we
                     have found that a small percentage, 2% to 3%, of the
```

```
ORIGIN
                     clones contain inserts that appear to be highly homologous
                     to sequences from salmonid fishes, as determined by
                     homology comparisons using BLAST and by Southern
                     hybridization analysis to genomic DNA from salmon (sigma
                     Chemical Co. St. Louis, MO) under stringent washing
                     conditions. The remainder of the clones appear to contain
                     S. japonicum sequences."
```

US-10-601-319-2 (1-440) x BU713770 (1-529)

```
Alignment Scores:
  Pred. No.:             1.08e-82        Length:          529
  Score:                 913.00          Matches:         174
  Percent Similarity:    98.9%           Conservative:    0
  Best Local Similarity: 98.9%           Mismatches:      2
  Query Match:           39.9%           Indels:          0
  DB:                    3               Gaps:            0
```

```
Qy   224  AlaAspAsnValSerLeuThrLeuGlyAlaValSerLeuAlaSerMetLeuThrGluIlePhe  243
Db     1  GCCGACAAATGTTCATTAAACCGGTGCGTGCGTAAAGCCTCGCATCAATGCTGACGGAGATATTT   60

Qy   244  LeuLeuGlnGlnAlaLaGlnGlyMetProGluProGlyTrpGlyArgIleThrAspSerHis  263
Db    61  CTCCTTGCCAACAAGCACAGGGAAATGCCGAGCCCGGGTGGGGGAGGATCACCGATTCACAC  120

Qy   264  GlnTrpAsnThrLeuLeuSerLeuHisAsnLeuGlyPheTyrIleLeuLeuGlnArgThrPro  283
Db   121  CAGTGGAACACCTTGCTAAGTTTGCATAACGGCGCAATTTTATTTGCTACAACGGCACGCCA  180

Qy   284  GluValAlaAlaArgSerArgAlaThrProLeuLeuMetAlaAlaAlaLeuThrPro  303
Db   181  GAGGTTGCCGCGCAGCCGGGCGCCACCCCGTTATTTAGATTTGATCAAGACAGCAGCGTTGACGCCC  240

Qy   304  HisProProGlnHisGlnAlaTyrGlyValThrHisLeuProThrSerValLeuPheIleAla  323
Db   241  CATCCACCGCAAAACAGGCGTATGGCGTGACATTACCCATCTTCAGTGGTCTGTTTATCGGCC  300

Qy   324  GlyHisLsAspThrArgAsnLeuAlaAsnLeuGlyValAlaLeuAsnTrpLeuPro  343
Db   301  GGACACGATACTAATTCGGCAAATCTCGGCGGGCGCCACTGGAGCTCAACTGGACCGCTTCCC  360

Qy   344  GlyGlnProAspAspAsnThrProGlyGlyGlyLeuValGluLeuValPheGluIleArgTrpProArgArgLeu  363
Db   361  GGTCAGCCCGGATAACACACCGGCCGGTGGTGAACTGGTGTTTGAACGCTGGCCGTCGGCTA  420

Qy   364  SerArgAspAsnSerGlnTrpIleGlnValSerLeuValPheGluLnThrArgLeuGlnIleMetArg  383
Db   421  AGCGGATAACAGCCAGTGGATTCAGGTTTCGCTGGTTCTTCCAGACTTTACAGCAGATGGGT  480

Qy   384  AspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeu  399
Db   481  GATAAAACGCCGCTGTCATTAAATACGGCCGCCCGGAGAGGGTGAAACTG  528
```

RESULT 4

```
Db         CA093060
LOCUS      CA093060              746 bp    mRNA    linear   EST 23-SEP-2003
DEFINITION SCCCCL2001D10.b CL2 Saccharum officinarum cDNA clone SCCCCL2001D10
           3', mRNA sequence.
ACCESSION  CA093060
VERSION    CA093060.1  GI:34946367
KEYWORDS   EST.
SOURCE     Saccharum officinarum
  ORGANISM Saccharum officinarum
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
           clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
           complex.
REFERENCE  1  (bases 1 to 746)
  AUTHORS  Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE    The libraries that made SUCEST
  JOURNAL  Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
```

```
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: Clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bccenter.fcav.unesp.br
            Plate: 001 row: D column: 10
            Seq primer: M13/Forward primer.
FEATURES             Location/Qualifiers
     source          1. .746
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCCCL2001D10"
                     /lab_host="XL1Blue MRF'"
                     /clone_lib="CL2"
                     /note="Organ: Pool of sugarcane calli submitted to low
                     (4oC) and high (37 C) temperature stress; Vector:
                     pBluescript Site 1: EcoRI; Site 2: XhoI; An
                     unidirectional cDNA library generated from [pool of
                     sugarcane calli submitted to low (4oC) and high (37 C)
                     temperature stress]. cDNA was prepared from polyA+ mRNA
                     using ZAP - cDNA Synthesis Kit (Stratagene). The
                     double-strand cDNAs were fractionated in a sepharose CL-2B
                     40cm-columns and fragments sizing between 0.8 and 1.5 Kb
                     were directionally cloned into the vector. Details of
                     each source of RNA and library construction can be
                     obtained at http://sucest.lad.ic.unicamp.br/public"
ORIGIN

US-10-601-319-2 (1-440) x CA093060 (1-746)

Alignment Scores:
Pred. No.:
Score:                   9.95e-81          Length:           746
Percent Similarity:      896.00            Matches:          177
Best Local Similarity:   96.7%             Conservative:     0
Query Match:             96.7%             Mismatches:       6
DB:                      38.9%             Indels:           1
                                           Gaps:             0

Qy  204 GluLysGlnAspGluSerCysSerSerLeuThrGlnAlaLeuProSerGluLeuLysValSer 223
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  199 GAGAAACAGGACGAGAGCTGTTCATTAACGGCAGGCATTACCATCGGAACTCAAGGTGAGC 258

Qy  224 AlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThrGluIlePhe 243
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  259 GCCGACAATGTCTCATTAACCGGTGCCGTAAGCCTCGCATCAATGCTGACGGAGATATT 318

Qy  244 LeuLeuGlnGlyGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThrAspSerHis 263
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  319 CTCCTGCAAGGACAAGCACAGGGAATGCCGGAGCCCGGGGTGGGGAGGATCACCGGATTCACAC 378

Qy  264 GlnTrpAsnThrLeuLeuLeuSerLeuHisAsnAlaGlnInPheTyrLeuLeuGlnArgThrPro 283
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  379 CAGTGGAACACCCTTGCTAAGTTTGCATAACCGGCAATTTATTTGCTACAACGCACGCA 438

Qy  284 GluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAlaLeuThrPro 303
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  439 GAGGGTTGCCCCGCAGCCCGCCACCCCGTTATTAGATTTGATCAAGACAGGTTGACGCCC 498

Qy  304 HisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeuPheIleAla 323
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  499 CATCCACCGACANAAACAGGCGGTTATGGTGTGACATTACCCACTTCAGTGCTGTTTTACGCC 558

Qy  324 GlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlyLeuAsnTrpIleProThrLeuPro 343
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  559 GGACACGGATACTAATCTGGCAAAATCCCGGCGGCACTGGGCGCTCAACTGGACGCTTCCC 618

Qy  344 GlyGlnProAspAsnThrProGlyGlyGlyGlyLeuValPheGluArgTrpArgArgLeu 363
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  619 GGTCAGCCGGATAACACGCCCCGCCAGGTAGTGGTTGAAACTGGTGTT-GAACGGCTGGGGTCNGCTA 677

Qy  364 SerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGlnIleMetArg 383
        ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||| |||
Db  678 AGCGATAACAGCCAGTCGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAGCANATGCGT 737

Qy  384 AspLysThr 386
        |||||||||
Db  738 GATAAAACG 746
```

```
RESULT 5
LOCUS       CF326092/c      595 bp    mRNA    linear    EST 18-AUG-2003
DEFINITION  JMT1--05-B09.g1 AtJMT-overexpressing transgenic rice lambda phage
            cDNA library (JMT1) Oryza sativa (japonica cultivar-group) cDNA
            clone JMT1--05-B09, mRNA sequence.
ACCESSION   CF326092
VERSION     CF326092.1  GI:33800445
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
ORGANISM    Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
            clade; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 595)
  AUTHORS   Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
            Song,S.I., Kim,J.K., Kim,J.K. and Nahm,B.H.
  TITLE     Large-scale Sequencing Analysis of Rice ESTs
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Nahm B.H.
            Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
            of Bioscience and Bioinformatics, Myongji University
            Yongin, Kyeonggi, Korea
            Tel: 82 31 330 6193
            Fax: 82 31 321 6355
            Email: bhnahm@mgbio.com, bhnahm@bio.myongji.ac.kr.
FEATURES             Location/Qualifiers
     source          1. .595
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nackdong"
                     /clone="JMT1--05-B09"
                     /db_xref="taxon:39947"
                     /clone_lib="AtJMT-overexpressing transgenic rice lambda
                     phage cDNA library (JMT1)"
                     /note="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2:
                     XhoI; cDNA was inserted into lamda Uni-ZAP XR vector at 5'
                     end with EcoRI and 3' end with XhoI site. mRNA was
                     prepared from Arabidopsis Jasmoinate Carboxyl
                     methyltransferase overexpression line."
ORIGIN

US-10-601-319-2 (1-440) x CF326092 (1-595)

Alignment Scores:
Pred. No.:
Score:                   9.25e-76          Length:           595
Percent Similarity:      846.00            Matches:          161
Best Local Similarity:   98.8%             Conservative:     0
Query Match:             98.8%             Mismatches:       2
DB:                      36.8%             Indels:           0
                                           Gaps:             0

Qy  270 SerLeuHisAsnAlaGlnInPheTyrLeuLeuGlnArgThrProGlyValAlaArgSerArg 289
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  595 AGTTTGCATAACCGCAATTTATTTGCTACAACGCACGCCAGAGGTTGCCCCGCAGCGCGC 536

Qy  290 AlaThrProLeuLeuAspLeuIleMetAlaAlaLeuThrProHisProGlnLysGln 309
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  535 GCCACCCCGTTATTAGATTTGATCAAGACAGGTTGACGCCCCATCCACCGACAAAAACAG 476
```

Left column

```
RESULT 6
LOCUS       CN762997/c
DEFINITION  ID0AAA5DC09RM1 APmS Acyrthosiphon pisum
            mRNA sequence.
ACCESSION   CN762997
VERSION     CN762997.1  GI:47536920
KEYWORDS    EST.
SOURCE      Acyrthosiphon pisum (pea aphid)
ORGANISM    Acyrthosiphon pisum
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
            Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
REFERENCE   1  (bases 1 to 895)
AUTHORS     Hunter,W., Martinez-Torres,D., Rahbe,Y., Sabater-Munoz,B.,
            Stern,D., Tagu,D. and Wincker,P.
TITLE       An expressed sequence tags database for the pea aphid Acyrthosiphon
            pisum
JOURNAL     Unpublished (2004)
COMMENT     Contact: D. Tagu
            INRA Rennes
            UMR BiO3P, BP 35327, F-35653 Le Rheu Cedex France
            Tel: +33.2.23.48.51.65
            Fax: +33.2.23.48.51.50
            Risk of contamination by bacterial sequences from obligatory
            (Buchnera) or facultative endosymbionts. These sequences were
            obtained in the frame of the International Consortium of Aphid
            Genomics in collaboration with Genoscope
PCR PRimers
FORWARD: CAGGAAACAGCTATGACC
Plate: 5   row: C   column: 9.
            Location/Qualifiers
1..895
/organism="Acyrthosiphon pisum"
/mol_type="mRNA"
/cultivar="developmentstage"
/db_xref="taxon:7029"
/clone="ID0AAA5DC09"
/tissue_type="whole insect"
/dev_stage="nymphs and adults (parthenogenetic females)"
/lab_host="XL1-Blue"
/clone_lib="APmS"
/note="Vector: pBS-SK minus; Site 1: EcoRI; Site 2: XhoI;
Sample name: ID0AAA ; Plant growth place: Department of
FEATURES
source
```

Sequence alignment (left):

```
310 AlaThrGlyValThrLeuProThrSerValLeuPheIleAlaGlyHisAspThrAsnLeu 329
475 GCCTATGGTGTGACATTACCCACTCACGTGTTATCGCCGGACACGATACTAATCTG
                                                              GG
GCAAATCTCGGCGCGGCGACGACTCAACTGGACGGCCTTCCCGGTGACGGATAACACG 356
415 AlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrpThrLeuProGlyGlnProAspAsnThr 349
330 IleGlnValSerLeuValPheGluIleGlnMetArgAspLysThrProLeuSer 389
370                                                         356
350 ProProGlyGlyLeuGluLeuValPheGlnIleAsnGluArgThrLeuAsn 369
355 CCCGCAGGTGGTGAACTGGTGTTTGAACGCTGGCGTGCGGATAACAGCCAGTGG 296
295 ATTCAGGTTTCGCTGGTCTTCGAAGTTCAGATGCGTGATAAAACGCCCGCTGTCA 236
390 LeuAsnThrProProGlyGlyLeuValGlyLysLeuAlaGlyLysCysGluGlyLeuuArgGlyAsnAla 409
235 TTAAATACGCCCCGTGGCGGCCCGGAGGTGAAAACTGACCCCTGGCGGGATGTGAAGAGCGAAAATCG 176
175 CAGGGCAATGTGGTTCGTTGGCAGGTTTTACGCAAAATCGTCAATGAAGACGCATACCCGGCG 116
430 CysSerLeu 432
115 TGCAGTTTG 107
```

Right column

```
US-10-601-319-2 (1-440) x CN762997 (1-895)

272 HisAlaAsnAlaGlnPheTyrLeuLeuGlnArgThrProGluValAlaAlaArgSerArgAlaThr 291
774 GGTGTGACATTACCCACTTCAGTCGTGTTTATCGCCGGACACGATACTAATCTGGCAAAT 715
894 CATAACGCCCAAATTTATTTGCTACAACCGCACGCCGGAGGTTGCCCGCCACCGCCCACC 835

292 ProLeuLeuAspLeuIleMetAlaAlaLeuuThrProHisProGlnLnLysGlnAlaTyr 311
834 CCGTTATTAGATTTGATTCAAGACAGCGTTGACGCCCCATCCACCGCAAAAACAGGCGTAT 775

312 GlyValThrLeuProThrSerValLeuPheIleAlaGlyHisAspThrAsnLeuuAlaAsn 331
834 GGTGTGACATTACCCACTTCAGTCGTGTTTATCGCCGGACACGATACTAATCTGGCAAAT 715

332 LeuGlyGlyAlaLeuGluLeuAsnTrpThrLeuProGlyGlnProAspAsnThrProPro 351
714 CTCGGCGCGGCACTGGAGCTCAACTGGACGGCCTTCCCGGTGACGGATAACACCCCGCCA 655

352 GlyGlyLeuGluLeuValPheGlnIleArgAspLysThrProLeuSerAspAsnSerGlnTrpIleGln 371
654 GGTGGTGAACTGGTGTTTGAACGCTGGCGTGCGGATGCGTGATAAAACGCCCGCTGTCATTAAAT 595

372 ValSerLeuValPheGluIleGlnMetArgAspLysThrProLeuSerAspAsn 391
594 GTTTCGCTGGTCTTCCAGACTTTACAGATGCGTGATAAAACGCCCGCTGTCATTAAAT 535

392 ThrProProGlyGlyLeuValGlyLysLeuAlaGlyLysCysGluGlyLeuuArgGlyAsnAlaGlnHisGlyGly 411
534 ACCGCCCGCCCGGAGAGGTGAAAACTGACCCCTGGCAGGTGTGAAGAGCGAAAATGGCAGGGC 475

412 MetCysSerLeuValAlaGlyPheThrGlnIleValAlaAsnGluAlaArgIleProAlaCysSer 431
474 ATGTGTTCGTTGGCAGGTTTTACGCAAAATCGTCAATGAAGACGCACGCCATACCCGGCGTGCAGT 415

432 Leu 432
414 TTG 412
```

```
RESULT 7
LOCUS       CN754382/c
DEFINITION  ID0AAA12DE01RM1 APmS Acyrthosiphon pisum cDNA clone ID0AAA12DE01
            5', mRNA sequence.
ACCESSION   CN754382
VERSION     CN754382.1  GI:47519379
KEYWORDS    EST.
SOURCE      Acyrthosiphon pisum (pea aphid)
ORGANISM    Acyrthosiphon pisum
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
            Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
REFERENCE   1  (bases 1 to 868)
AUTHORS     Hunter,W., Martinez-Torres,D., Rahbe,Y., Sabater-Munoz,B.,
            Stern,D., Tagu,D. and Wincker,P.
TITLE       An expressed sequence tags database for the pea aphid Acyrthosiphon
```

```
  JOURNAL   Unpublished (2004)
  COMMENT   Contact: D. Tagu
            INRA Rennes
            UMR BiO3P, BP 35327, F-35653 Le Rheu Cedex France
            Tel: +33.2.23.48.51.65
            Fax: +33.2.23.48.51.50
            Risk of contamination by bacterial sequences from obligatory
            (Buchnera) or facultative endosymbionts. These sequences were
            obtained in the frame of the International Consortium of Aphid
            Genomics in collaboration with Genoscope
            PCR PRimers
            FORWARD: CAGGAAACAGCTATGACC
     Plate: 12 row: E column: 1.
  FEATURES  Location/Qualifiers
     source   1..868
            /organism="Acyrthosiphon pisum"
            /mol_type="mRNA"
            /cultivar="developmentstage"
            /db_xref="taxon:7029"
            /clone="ID0AAA12DE01"
            /tissue_type="whole insect"
            /dev_stage="nymphs and adults (parthenogenetic females)"
            /lab_host="Xl1-Blue"
            /clone_lib="ApMS"
            /note="Vector: pBS-SK minus; Site_1: EcoRI; Site_2: XhoI;
            Sample name: ID0AAA ; Plant growth place: Department of
            Ecology & Evolutionary Biology, Princeton University ;
            Soil conditions: Soil ; Sowing date: 01/06/1999 ;
            Harvesting date: 01/06/1999 ; Stress date: no stress ;
            Description: Aphids inoculated on one-week old Vicia faba
            under non-sterile conditions. All parthenogenetic stages
            and both winged and wingless adults were collected for
            library construction. ; experimental condition: long
            photoperiod (16-hr light/8-hr dark at 18 c)"
  ORIGIN

US-10-601-319-2 (1-440) x CN754382 (1-868)

Alignment Scores:
Pred. No.:              1.99e-67            Length:        868
Score:                  767.00             Matches:       149
Percent Similarity:     98.0%              Conservative:  0
Best Local Similarity:  98.0%              Mismatches:    3
Query Match:            33.3%              Indels:        1
DB:                     8                  Gaps:          0

Qy  281 ArgThrProGlyuValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAlaAla 300
            ||||||||||||| |||||||||||||||||||||||||||||||||||||||| |||
Db  866 CGCACGCCAGAGTT-GCCCGCGCAGCCCACCGGCCTTATTAGATTTGATCAAGACAGCG 808

Qy  301 LeuThrProHisProProGlnLysGlnAlaThrTyrGlyValThrLeuProThrSerValLeu 320
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  807 TTGACCCCCATCCACCGCAAAAACAGGCGTANGGTGTGACATTACCACTTCAGTGCTG 748

Qy  321 PheIleAlaGlyHisAspThrAsnLeuuAlaAsnLeuGlyGlyAlaLeuGlyLeuAsnTrp 340
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  747 TTTATCGCCCGGACACCGATACTAAATCTGGCAAAATCTCGGCGCCCGCCACTCAACTGG 688

Qy  341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  687 ACGGCTTCCCGGTCCAGCCGGATAACACGCCCCCAGGTGGTGAACTGGTGTTGAACCGTGG 628

Qy  361 ArgArgLeuSerAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrrLeuGln 380
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  627 CGTCGGCTAAGCGATAACAGCCAGGTTGGATTCAGGTTTCGCTGCTCTTCCAGACTTTACAG 568

Qy  381 GlnMetArgAsnpLysThrProLeuSerLeuAsnThrProProGlyGluVaIlLysLeuThr 400
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  567 CAGAGCGGTGATAAAAACGCCCCTCGTCATTAAATACCCCGCCCGGAGAGGTGAAAACTGACC 508

Qy  401 LeuAlaGlyCysGluGluuArgAsnAlaGlnGlyMetCysSerSerLeuAlaGlyPheThrGln 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  507 CTGGCAGGATGTGAGAGCGAAAATGGGCAGGGCATGTGTTCGTTGGTTGGCAGGTTTTACGGCAA 448

Qy  421 IleValAsnGluAlaArgIleProAlaCysSerSerLeu 432
            |||||||||||||||||||||||||||||||||||
Db  447 ATCGTGAATGAAGCACGCCATACCGGCGGTGCAGTTTG 412
```

RESULT 8

```
  LOCUS       CN757565              864 bp    mRNA    linear    EST 20-MAY-2004
  DEFINITION  ID0AAA1DF09RM1 ApMS Acyrthosiphon pisum cDNA clone ID0AAA1DF09 5',
              mRNA sequence.
  ACCESSION   CN757565
  VERSION     CN757565.1  GI:47531488
  KEYWORDS    EST.
  SOURCE      Acyrthosiphon pisum (pea aphid)
    ORGANISM  Acyrthosiphon pisum
              Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
              Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidoidea;
              Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
  REFERENCE   1  (bases 1 to 864)
    AUTHORS   Hunter,W., Martinez-Torres,D., Rahbe,Y., Sabater-Munoz,B.,
              Stern,D., Tagu,D. and Wincker,P.
    TITLE     An expressed sequence tags database for the pea aphid Acyrthosiphon
              pisum
    JOURNAL   Unpublished (2004)
  COMMENT     Contact: D. Tagu
              INRA Rennes
              UMR BiO3P, BP 35327, F-35653 Le Rheu Cedex France
              Tel: +33.2.23.48.51.65
              Fax: +33.2.23.48.51.50
              Risk of contamination by bacterial sequences from obligatory
              (Buchnera) or facultative endosymbionts. These sequences were
              obtained in the frame of the International Consortium of Aphid
              Genomics in collaboration with Genoscope
              PCR PRimers
              FORWARD: CAGGAAACAGCTATGACC
       Plate: 1 row: F column: 9.
  FEATURES    Location/Qualifiers
       source   1..864
              /organism="Acyrthosiphon pisum"
              /mol_type="mRNA"
              /cultivar="developmentstage"
              /db_xref="taxon:7029"
              /clone="ID0AAA1DF09"
              /tissue_type="whole insect"
              /dev_stage="nymphs and adults (parthenogenetic females)"
              /lab_host="Xl1-Blue"
              /clone_lib="ApMS"
              /note="Vector: pBS-SK minus; Site_1: EcoRI; Site_2: XhoI;
              Sample name: ID0AAA ; Plant growth place: Department of
              Ecology & Evolutionary Biology, Princeton University ;
              Soil conditions: Soil ; Sowing date: 01/06/1999 ;
              Harvesting date: 01/06/1999 ; Stress date: no stress ;
              Description: Aphids inoculated on one-week old Vicia faba
              under non-sterile conditions. All parthenogenetic stages
              and both winged and wingless adults were collected for
              library construction. ; experimental condition: long
              photoperiod (16-hr light/8-hr dark at 18 c)"
  ORIGIN

US-10-601-319-2 (1-440) x CN757565 (1-864)

Alignment Scores:
Pred. No.:              2.09e-66            Length:        864
Score:                  757.00             Matches:       148
Percent Similarity:     98.0%              Conservative:  0
Best Local Similarity:  98.0%              Mismatches:    3
Query Match:            32.9%              Indels:        1
DB:                     8                  Gaps:          0

Qy  282 ThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAlaLeu 301
            |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db  863 ACGCCAGAGGTTGCCCGCGCAGCCCACCCGTTATTAGATTG-ATCAAGACAGCAGCGTTG 805
```

RESULT 9

```
AW036132
LOCUS       AW036132       383 bp    mRNA    linear    EST 18-MAY-2001
DEFINITION  EST274508 tomato seed, TAMU Lycopersicon esculentum cDNA clone
            cLEE1E23 similar to periplasmic phosphoanhydride phosphohydrolase
            precursor, putative, mRNA sequence.
ACCESSION   AW036132
VERSION     AW036132.1  GI:5894811
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM    Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 383)
AUTHORS     Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,
            Holt,I.E., Liang,F., Upton,J., Ronning,C.M., Craven,C.M.,
            Fujii,C.Y., Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C.,
            Martin,G.B., Tanksley,S.D. and Giovannoni,J.
TITLE       Generation of ESTs from tomato seed tissue
JOURNAL     Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
FEATURES    Location/Qualifiers
    source  1..383
            /organism="Lycopersicon esculentum"
            /mol_type="mRNA"
            /cultivar="TA496"
            /db_xref="taxon:4081"
            /clone="cLEE1E23"
            /tissue_type="seeds"
            /dev_stage="quiescent seed"
            /lab_host="XL1-Blue MRF'"
            /clone_lib="tomato seed, TAMU"
            /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
            XhoI; cLEE - Tomato Seed EST Library. Directionally cloned
            cDNAs inserted into pBluescript SK(-) at 5' end and with
            EcoRI and 3' end with XhoI site."
ORIGIN
```

```
Qy  302  ThrProHisProProGlnLysGlnAlaThrGlyValThrLeuProThrSerValLeuPhe  321
Db  804  ACGCCCATCCACCGAAAAACAGGCGGTATGGGTGTGACATTACCCACTTCAGTGCTGTTT  745

Qy  322  IleAlaGlyHisAlaAspThrAsnLeuAlaAsnLeuGlyLysGlyAlaLeuGluTrpThr  341
Db  744  ATCGCCGGACACGATAATCTGGCCAATCTCGGCCGGCGCACTGGAGCTCAACTGGACG    685

Qy  342  LeuProIleProGlnAsnThrProProGlyGlyGlyLeuValPheGluGlnArgTrpArg  361
Db  684  CTTCCCCGGTCAGCCGGATAACACGCCCCCAGGTGGTGGAACTCGTGTTTGAACCCTGGT  625

Qy  362  ArgLeuSerAspAsnSerGlnThrTrpIleGlnIleValSerLeuValPheGlnThrLeuGlnGln  381
Db  624  CGGCTAAGCGATAACAGCCAGGTTTGGTAGATTGTGTCTCTGGTCTTTCCAGACTTTACAGGCAG  565

Qy  382  MetArgAspGlyThrProLeuSerLeuAsnThrProProGlyGlyGluValLysGlyLeuThrLeu  401
Db  564  ATGCGTGATGGAAACGCCCTGTCATTAAATACGCCCCGCGGGGAGGTGAAAACTGACCCTG     505

Qy  402  AlaGlyCysGluGluGlnArgAspAsnAlaGlnGlyMetCysSerLeuValPheGlyThrGlnIle  421
Db  504  GCAGGATGTGAAGAGCCAAGAATCGCAGGGAATGTGTTCGTTGGCAGGTTTTACGGCAAATC      445

Qy  422  ValAsnGluIleLeuArgIleProAlaCysSerSerLeu  432
Db  444  GTGAATGAAGCACGCATACCGGCGTCGCAGTTTG      412
```

Alignment Scores:

Pred. No.:	2.83e-55	Length:	383
Score:	643.00	Matches:	127
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	27.9%	Indels:	0
DB:		Gaps:	0

US-10-601-319-2 (1-440) x AW036132 (1-383)

RESULT 10

```
AW036134/c
LOCUS       AW036134       383 bp    mRNA    linear    EST 18-MAY-2001
DEFINITION  EST274510 tomato seed, TAMU Lycopersicon esculentum cDNA clone
            cLEE1E23 similar to periplasmic phosphoanhydride phosphohydrolase
            precursor, putative, mRNA sequence.
ACCESSION   AW036134
VERSION     AW036134.1  GI:5894813
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM    Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 383)
AUTHORS     Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,
            Holt,I.E., Liang,F., Upton,J., Ronning,C.M., Craven,C.M.,
            Fujii,C.Y., Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C.,
            Martin,G.B., Tanksley,S.D. and Giovannoni,J.
TITLE       Generation of ESTs from tomato seed tissue
JOURNAL     Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
FEATURES    Location/Qualifiers
    source  1..383
            /organism="Lycopersicon esculentum"
            /mol_type="mRNA"
            /cultivar="TA496"
            /db_xref="taxon:4081"
            /clone="cLEE1E23"
            /tissue_type="seeds"
5 prime sequence.
```

```
Qy  117  LysThrGlyGluAlaPheAlaAlaGlyLeuAlaAlaProAspProCysAlaIleThrValHisThr  136
Db    2  AAAACAGGCGGAAGCCTTCGCCGCCGGGCTGGCAGCACCTGACTGTGCAATAACGTACATACC    61

Qy  137  GlnAlaAspThrSerSerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGln  156
Db   62  CAGGCAGATACGTCCCCGATCCGTTATTTAATCCTCTAAAAACTGGGGTTTGCCAA       121

Qy  157  LeuAspAsnAlaAsnValThrArgAspAlaIleLeuGluSerArgAlaGlyGlySerIleAlaAsp  176
Db  122  CTGGATAACGCGAACGTGACACTGACGGCAATCCTCAGCAGGGCGGGAGGTTCAATGGCTGAC    181

Qy  177  PheThrArgGlyHisArgGlnThrAlaPheAlaArgGluGluLeuGluValValLeuAsnPheProGln  196
Db  182  TTTACCCGGGGCCATCGGCAAACGGGCGTTTCCGGAACTGGAAGAGGGGTTGCTTAAATTTCCGCAA  241

Qy  197  SerAsnLeuCysSerLeuValArgGluGluLysGlnAspProAspSerCysSerLeuThrGlnIleAlaLeu  216
Db  242  TCAAACTTGTGCCTTAAACGTGAGAAAACAGGACGAAAAGCTGTTCATTAACGTCAGGCAGCATTA  301

Qy  217  ProSerGluLeuLysValSerSerAlaAlaAspAsnAlaValSerLeuThrGlyAlaValValSerLeuValAla  236
Db  302  CCATCCGGAACTCAAGGTGAGCGCCGGACAATGTCTCATTAACCCGGTGTGCGGTAAGCCTCGGCA    361

Qy  237  SerMetLeuThrGluIlePhe  243
Db  362  TCAATGCTGACCGGAGATATT   382
```

```
FEATURES
     source          1..393
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /ecotype="Columbia"
                     /db_xref="taxon:3702"
                     /clone="M11B12"
                     /tissue_type="seed"
                     /dev_stage="5-13 days after flowering"
                     /lab_host="E.coli"
                     /clone_lib="Arabidopsis developing seed"
                     /note="Organ: Developing seed; Vector: pBluescript SK-;
                     Site_1: EcoRI; Site_2: XhoII"

                     /dev_stage="quiescent seed"
                     /lab_host="X11-Blue MRF'"
                     /clone_lib="tomato seed, TAMU"
                     /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
                     Xhol; cLEE - Tomato Seed EST Library. Directionally cloned
                     cDNAs inserted into pBluescript SK(-) at 5' end with
                     EcoRI and 3' end with Xhol site."

ORIGIN

US-10-601-319-2 (1-440) x AW036134 (1-383)

Alignment Scores:
Pred. No.:
Score:                 2.83e-55
Score:                 643.00         Length:         383
Percent Similarity:    100.0%         Matches:        383
Best Local Similarity: 100.0%         Conservative:   127
Query Match:           27.9%          Mismatches:     0
DB:                                   Indels:         0
                                      Gaps:           0

Qy  117 LysThrGlyGluAlaAlaPheAlaAlaGlyLeuLeuAlaProAspCysAlaIleThrValHisThr 136
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  382 AAAACAGGCGGAAGCCTTCGCCGCGCCTGGCACTTGACGTGTGCAATAACCGTACNATACC 323

Qy  137 GlnAlaAspThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGln 156
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  322 CAGGCGAGATACGTCCAGTCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAA 263

Qy  157 LeuAspAsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAsp 176
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  262 CTGGATAACGCGAACGTGACTGACGCGATCCTCAGCCAGGGGCAGGAGGGTCAATTGCTGAC 203

Qy  177 PheThrGlyHisArgGlnThrArgAlaPheArgGluLeuGluArgValLeuAsnProGln 196
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  202 TTTACCGGGCATCGGCAAACCGCGTTTCGCGAACTGGAACGGGTGCTTAATTTCCGGCAA 143

Qy  197 SerAsnLeuCysLeuLysArgGlyValGlyLysGlnSerCysSerLeuThrGlnAlaLeu 216
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  142 TCAAACTTGTGCCTTAAACGTGGAGAAACAGAGCTGTTCATTAACGCAGGCATTA 83

Qy  217 ProSerGluLeuLysValSerAlaAspAsnValSerLeuThrGlyValAlaValSerLeuAla 236
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  82 CCATCGGAACTCAAGGTGAGCGCCGACAACAATCTCATTAACGGTGCGGTAAGCCTCGCA 23

Qy  237 SerMetLeuThrGlyIlePhe 243
        |||||||||||||||||||||
Db  22 TCAATGCTGACGGAGATATTT 2
```

RESULT 11
BES20240/c

```
LOCUS       BES20240           393 bp    mRNA    linear   EST 19-MAR-2001
DEFINITION  M11B12STM Arabidopsis developing seed Arabidopsis thaliana cDNA
            clone M11B12 5', mRNA sequence.
ACCESSION   BES20240
VERSION     BES20240.1  GI:9778242
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 393)
AUTHORS     White,J.A., Todd,J., Newman,T., Focks,N., Girke,T., Martinez de
            Ilarduya,O., Jaworski,J.G., Ohlrogge,J. and Benning,C.
TITLE       A new set of Arabidopsis expressed sequence tags from developing
            seeds. The metabolic pathway from carbohydrates to seed oil
JOURNAL     Plant Physiol. 124 (4), 1582-1594 (2000)
PUBMED      11115876
COMMENT     Contact: Benning, C
            Dept. of Biochemistry & Molecular Biology
            Michigan State University
            224 Biochemistry, Michigan State University, East Lansing, MI
            48824, USA
            Tel: 517 355 1609
            Fax: 517 353 9334
            Email: benning@msu.edu
            Michigan State University DNA Sequencing Facility Arabidopsis
            Biological Resource Center, The Ohio State University, 309 Botany &
            Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX:
            61429206 03 TEL: 614292937.
```

ORIGIN

US-10-601-319-2 (1-440) x BES20240 (1-393)

```
Alignment Scores:
Pred. No.:
Score:                 4.71e-55
Score:                 641.00         Length:         393
Percent Similarity:    97.7%          Matches:        393
Best Local Similarity: 96.9%          Conservative:   127
Query Match:           27.8%          Mismatches:     1
DB:                                   Indels:         3
                                      Gaps:           0

Qy  107 IleIleAlaAspValAspGluArgThrArgLysThrGlyGluAlaPheAlaAlaGlyLeu 126
        ||||||||||||||| |||||||||||:|||||||||||||||||||||||||||||||
Db  393 ATTATTGCTGATGTCGACGAGCGTACCCGTANACAGGGCGAAGCTTTCGCCGCCGGGNTG 334

Qy  127 AlaProAspCysAlaIleThrValHisThrGlnAlaAspThrSerSerProAspProLeu 146
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  333 GCAACCTGACTGCAATAACCGTACATACCCAGGCAGATACGTCCGATCCGTTA 274

Qy  147 PheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAlaAsnValThrAspAlaIle 166
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  273 TTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCGAACGTGACGGGATC 214

Qy  167 LeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHisArgGlnThrArgAlaPheArg 186
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  213 CTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCATCGGCAAACCGGCGTTTCGC 154

Qy  187 GluLeuGluArgValLeuAsnProGlnSerAsnLeuCysLeuLysArgGlyLysGln 206
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db  153 GAACTGGAACGGGTGCTTAATTTTCCGCAACTAAACTTGTGCCTTAAACGTGGAGAAACAG 94

Qy  207 AspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeuLysValSerAlaAspAsn 226
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  93 GACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTCAAGGTGAGCGCCGACAAT 34

Qy  227 ValSerLeuThrGlyAlaValAlaValSerLeuAlaSer 237
        |||||||||||:||||||||||||||||||||||
Db  33 GTCTCATTAACGGCTGCGGTAAGCCTCGCATCA 1
```

RESULT 12
BU713769

```
LOCUS       BU713769           531 bp    mRNA    linear   EST 23-OCT-2003
DEFINITION  SJAABUG01 Adult SjC 7/94 Schistosoma japonicum cDNA similar to
            sp|P07102|PPA_ECOLI PERIPLASMIC APPA PROTEIN PRECURSOR [INCLUDES:
            PHOSPHOANHYDRIDE PHOSPHOHYDROLASE (PH 2.5 ACID PHOSPHATASE) (AP);
            6-PHYTASE ], mRNA sequence.
ACCESSION   BU713769
VERSION     BU713769.1  GI:28321125
KEYWORDS    EST.
SOURCE      Schistosoma japonicum
ORGANISM    Schistosoma japonicum
            Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
            Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
```

REFERENCE   1 (bases 1 to 531)
AUTHORS   Hu,W., Yan,Q., Shen,D.K., Liu,F., Zhu,Z.D., Song,H.D., Xu,X.R.,
             Wang,Z.J., Rong,Y.P., Zeng,L.C., Wu,J., Zhang,X., Wang,J.J.,
             Xu,X.N., Wang,S.Y., Fu,G., Zhang,X.L., Wang,Z.Q., Brindley,P.J.,
             McManus,D.P., Xue,C.L., Feng,Z., Chen,Z. and Han,Z.G.
TITLE   Evolutionary and biomedical implications of a Schistosoma japonicum
             complementary DNA resource
JOURNAL   Nat. Genet. 35 (2), 139-147 (2003)
PUBMED   12973349
COMMENT   Contact: Zeguang Han
             Chinese National Human Genome Center at Shanghai
             351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
             201203, P. R. China
             Tel: 86-21-50801919(ex.45)
             Fax: 86-21-50801922
             Email: hanzg@chgc.sh.cn.
FEATURES   Location/Qualifiers
   source   1..531
             /organism="Schistosoma japonicum"
             /mol_type="mRNA"
             /strain="Chinese (Anhui) strain"
             /db_xref="taxon:6182"
             /sex="Male and female"
             /tissue_type="Whole body"
             /dev_stage="Adult worms"
             /lab_host="Mouse and rabbit"
             /clone_lib="Adult SjC 7/94"
             /note="Vector: Lambda ZAP-II XR.; Site_1: EcoR I; Site_2:
             Xho I; Several hundred adult Schistosoma japonicum
             from the mesenteries of experimentally infected mice and
             rabbits (Anhui, P.R. China, strain), of mixed sex, were perfused
             Brisbane, Australia (QIMR), and stored for several months
             in liquid nitrogen. Subsequently, mRNA was isolated at the
             QIMR from lysates of these worms by oligo dT
             chromatography, using a kit from Pharmacia. The mRNA was
             then shipped to Clontech, Palo Alto, CA, USA, who
             constructed a cDNA library. First strand synthesis was
             primed with an oligo-dT-XhoI-primer and synthesized using
             M-MLV reverse transcriptase. Second strand synthesis was
             accomplished with RNAse H and T4 DNA polymerase. The
             double stranded cDNA was ligated to EcoRI linkers,
             digested with EcoRI and XhoI, and ligated into the
             phagemid vector lambda ZAP II XR. After construction of
             this directional library by Clontech, it was returned to
             the QIMR. During analysis of the library at the QIMR, we
             have found that a small percentage, 2% to 3%, of the
             clones contain inserts that appear to be highly homologous
             to sequences from salmonoid fishes, as determined by
             homology comparisons using BLAST and by Southern
             hybridization analysis to genomic DNA from salmon (Sigma
             Chemical Co., St. Louis, MO) under stringent washing
             conditions. The remainder of the clones appear to contain
             S. japonicum sequences."

ORIGIN

Alignment Scores:
Pred. No.:   8.65e-50
Score: 591.50
Percent Similarity: 79.5%   Length: 531
Best Local Similarity: 75.6%   Matches: 133
Query Match: 25.7%   Conservative: 7
                                    Mismatches: 30
                                    Indels: 7
                                    Gaps: 1

US-10-601-319-2 (1-440) x BU713769 (1-531)

Qy 224 AlaAspProAsnValSerLeuThrGlyAlaValAlaSerLeuHisSerMetLeuThrGluIlePhe 243
Db   1 GCCGACAAATGTCTCATTAACCGGTGCCGTAAGCCTCGCATCAATGCTGACCGGAGATCGGG 60

Qy 244 LeuLeuGlnGlnAlaGlnGlyMetProGlnProGlyTyrGlyArgIleThrAspSerHis 263
Db   61 CTCCTGCAACAACGCACAGGGAATGCCGAGCCGGGGTGGGGGGAAGGATCACCGGATTCACAC 120

Qy 264 GlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGlnArgThrPro 283
Db 121 CAGTGGAACACCTTGCTAAGTTTGCATAAGCGCCAAATTTATTTGCTACAACGCACGCCA 180

Qy 284 GluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAlaAlaThrPro 303
Db 181 GAGGTTGCCCGCAGCCGCGCCACCCCGGTATTAGATTTGATCAAGGACAGGTTGACGCGCC 240

Qy 304 HisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu-PheIleAl 323
Db 241 CGTTTATCATTATGATTTGCATGTGGAGAGACAATTCCCTTCTTGAGTGAGCGAAAAAAGG 300

Qy 323 aGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrpThrLeuPr 343
Db 301 CGGGCCACGAGCCTGAGCTGATAAAAAACTCGGCGCGGGCGGGCCACTGGGACCTGGCGGT 360

Qy 343 oGlyGlnProAspAsnThrProProGlyGlyGluLeu-ValPheGluAsnGlyTrp-ArgArg 362
Db 361 CAGGCAGCCCCGATAACACCCCCGAGGAACTGGATGTTTGAACCGCTGGGGCGTCGG 420

Qy 363 LeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlyGlnThrLeuGlnMet 382
Db 421 CTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTTCTTCCAGACTTA-CAGGAGATG 479

Qy 383 ArgAspLysThrProLeuSerLeuAsnThrProProGlyGlu 396
Db 480 CGTCGGGGGGGGTTCCTTTGGGG------CCTTATGGGGAT 512

RESULT 13
CN759004/c
LOCUS   CN759004   706 bp   mRNA   linear   EST 20-MAY-2004
DEFINITION ID0AAA24BC04RM1 ApMS Acyrthosiphon pisum cDNA clone ID0AAA24BC04
            5', mRNA sequence.
ACCESSION   CN759004
VERSION   CN759004.1 GI:47532927
KEYWORDS   EST.
SOURCE   Acyrthosiphon pisum (pea aphid)
  ORGANISM Acyrthosiphon pisum
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
            Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
REFERENCE   1 (bases 1 to 706)
AUTHORS   Hunter,W., Martinez-Torres,D., Rahbe,Y., Sabater-Munoz,B.,
            Stern,D., Tagu,D. and Wincker,P.
TITLE   An expressed sequence tags database for the pea aphid Acyrthosiphon
            pisum
JOURNAL   Unpublished (2004)
COMMENT   Contact: D. Tagu
            INRA Rennes
            UMR BiO3P, BP 35327, F-35653 Le Rheu Cedex France
            Tel: +33.2.23.48.51.55
            Fax: +33.2.23.48.51.50
            Risk of contamination by bacterial sequences from obligatory
            (Buchnera) or facultative endosymbionts. These sequences were
            obtained in the frame of the International Consortium of Aphid
            Genomics in collaboration with Genoscope
            PCR Primers
            FORWARD: CAGGAAACAGCTATGACC
            Plate: 24 row: C column: 4.
FEATURES   Location/Qualifiers
   source   1..706
            /organism="Acyrthosiphon pisum"
            /mol_type="mRNA"
            /cultivar="developmental stage"
            /db_xref="taxon:7029"
            /clone="ID0AAA24BC04"
            /tissue_type="whole insect"
            /dev_stage="nymphs and adults (parthenogenetic females)"
            /clone_lib="ApMS"
            /lab_host="XL1-Blue"
            /note="Vector: pBS-SK minus; Site_1: EcoRI; Site_2: XhoI;
            Sample name: ID0AAA ; Plant growth place: Department of

ORIGIN

Ecology & Evolutionary Biology, Princeton University ;
Soil conditions: Soil ; Sowing date: 01/06/1999 ; Stress date: no stress ;
Harvesting date: 01/06/1999 ; Stress date: no stress
Description: Aphids inoculated on one-week old Vicia faba
under non-sterile conditions. All parthenogenetic stages
and both winged and wingless adults were collected for
library construction. ; experimental condition: long
photoperiod (16-hr light/8-hr dark at 18 c)"
/clone="HBMSF1B4"
/sex="Male and Female"
/tissue_type="bone marrow"
/cell_type="stromal fibroblast"
/dev_stage="mixed"
/lab_host="XLl-Blue"
/clone_lib="Human Bone Marrow Stromal Fibroblast"
/note="Vector: pBluescript; Site_1: EcoRI; Site_2: XhoI"

US-10-601-319-2 (1-440) x CN759004 (1-706)

Alignment Scores:
Pred. No.:   1.73e-42
Score: 522.00
Percent Similarity: 100.0%   Length: 706
Best Local Similarity: 100.0%   Matches: 98
Query Match: 22.7%   Conservative: 0
DB: 8   Mismatches: 0
                                  Indels: 0
                                  Gaps: 0

Qy 335 AlaLeuGluLeuAsnTrpThrLeuProGlyGlnProAlaProGlyGlyGlu 354
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   705 GCACTGGAGCTCAACTGGACACTTCCCGGTCAGCCCGATAACACGCCGGTGAA 646

Qy 355 LeuValPheGluArgTrpArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeu 374
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   645 CTGGTGTTTGAACGCTGGCGTCTGAGCGATAACAGCCAGTGGATTCAGGTTTCGCTG 586

Qy 375 ValPheGlnThrLeuGlnGlnMetArgAspLysThrProLeuSerLeuAsnThrProPro 394
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   585 GTCTTCCAGACMTTTACAGCAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCC 526

Qy 395 GlyGluValLysGluThrLeuAlaGlyCysGluGlyLysArgAsnAlaGlnGlyMetCysSer 414
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   525 GGAGAGGTGAAAACTGACCCTGGCGACGATGTGAAGAGGCGAAATGCCCAGGGCATGTGTTCG 466

Qy 415 LeuAlaGlyPheThrGlnIleValAsnGluAlaArgIleProAlaCysSerLeu 432
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   465 TTGGCAGGTTTTACGCAAATCGTGAATGAAGCGCAATACCCGGCGTGCAGTTG 412

RESULT 14
AA545747/c
LOCUS   AA545747   354   mRNA   linear   EST 12-MAY-1999
DEFINITION HBMSF1B4-REV Human Bone Marrow Stromal Fibroblast Homo sapiens cDNA
            clone HBMSF1B4 5', mRNA sequence.
ACCESSION   AA545747
VERSION   AA545747.1 GI:2307026
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1 (bases 1 to 354)
AUTHORS   Jia,L., Robey,P., Young,M., Schuler,G., Powell,J., Yang,L.,
            Lennon,G., Hillier,L., Allen,M., Bowles,L., Geisel,S., Kucaba,T.,
            Marra,M., Martin,J., Steptoe,M., Wilson,R. and Francemano,C.
            Wylie,T., Waterston,R., Wilson,R. and Francemano,C.
TITLE   WashU-MGB/NHGRI EST Project
JOURNAL   Unpublished (1997)
COMMENT   Contact: Libin Jia
            Medical Genetics Branch
            National Human Genome Research Institute
            10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
            Tel: 301-402-4877
            Fax: 301-496-7157
            Email: libin@helix.nih.gov
            Seq primer: M13 Reverse
FEATURES   Location/Qualifiers
source   1..354
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"

ORIGIN

US-10-601-319-2 (1-440) x AA545747 (1-354)

Alignment Scores:
Pred. No.:   7.2e-42
Score: 511.50
Percent Similarity: 95.0%   Length: 354
Best Local Similarity: 94.1%   Matches: 112
Query Match: 22.2%   Conservative: 1
DB: 1   Mismatches: 5
                                  Indels: 4
                                  Gaps: 0

Qy 166 IleLeuSerArgAlaGlyGlySerIleAlaLeuAlaAspPheThrGlyHisArgGlnThr-AlaPh 185
         ||: ||||||||||||||||||||||||||||||||||||||||||
Db   352 GTCTTCAGCAGGGCAGGAGGTCA-ATTGCTGA-TTTACCGGCGCCATCGGCAAACGGGCCGGTT 295

Qy 185 eArgGluLeuGluArgValLeuAsnPheProGlnSerArgAsnLeuCysLeuLysArgGluLy 205
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   294 TCGCGGAACTGGAACGGGTGGTT-AATTTTCCGCAATCAAAACTTGTGCCTTAAACGTGAGAA 236

Qy 205 sGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeuLysValSerAlaAs 225
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   235 ACAGGACGGAAAACGTGTCATTAACGCAGGCATTACCATCGGGAACTCAAGGTGAGCGCCGCGA 176

Qy 225 pAsnValSerLeuThrGlyAlaAlaValSerLeuLeuAlaSerMetLeuThrGluIlePheLeuLe 245
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   175 CAATGNTNTCATTAACCCGTGCGGTAACCCTCCCATCAATGCTGACCGAGATATTTNTCCT 116

Qy 245 uGlnGlnAlaGlnGlyMetProGluProGlyThrGlyTrpGlyArgIleThrArgAspSerHisGlnTr 265
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   115 GCAACAGGCACAGGGAATGCCCGAGCCCGGGACCGGGGATCACCGGATTCACACCACCAGTG 56

Qy 265 pAsnThrThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGlnArgThrPro 283
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   55 GAACACCTTGCTAAGTTTGCATAAACGGCAATTTTATTTGCTACAACGCACGCCA 1

RESULT 15
BJ619443
LOCUS   BJ619443   707   mRNA   linear   EST 01-OCT-2003
DEFINITION BJ619443 NIBB Mochii normalized Xenopus early gastrula library
            Xenopus laevis cDNA clone XLl89d24 5', mRNA sequence.
ACCESSION   BJ619443
VERSION   BJ619443.1 GI:37258203
KEYWORDS   EST.
SOURCE   Xenopus laevis (African clawed frog)
ORGANISM   Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Xenopus.
REFERENCE   1 (bases 1 to 707)
AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and
            Kohara,Y.
TITLE   Expressed genes in X. laevis embryo
JOURNAL   Unpublished (2001)
COMMENT   Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshin@genes.nig.ac.jp
            The information of this clone is available through the following
            URL.
            http://xenopus.nibb.ac.jp.
FEATURES   Location/Qualifiers

ORIGIN

source   1. .707
                /organism="Xenopus laevis"
                /mol_type="mRNA"
                /db_xref="taxon:8355"
                /clone="XL189d24"
                /tissue_type="whole embryo"
                /dev_stage="stage 10.5"
                /clone_lib="NIBB Mochii normalized Xenopus early gastrula
                library"

Alignment Scores:
Pred. No.:   Length: 707
Score: 2.66e-34   Matches: 83
Percent Similarity: 97.6%   Conservative: 0
Best Local Similarity: 97.6%   Mismatches: 2
Query Match: 19.2%   Indels: 0
DB: 2   Gaps: 0

US-10-601-319-2 (1-440) x BJ619443 (1-707)

Qy 251 MetProGluProGlyTrpGlyTyrGlyArgIleThrArgSerHisGlnTrpAsnThrLeuLeuSer 270
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2 ATGCCCGAGCCCGGGGTGGGGGAAGGATCACCGAATCCACACCAGTGGGACACCTTGCTAAGT 61

Qy 271 LeuHisAsnAlaGlnPheTyrLeuLeuGlnArgThrProGluValAlaAlaArgSerArgAla 290
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62 TTGCATAACGCGCAATTTTATTTGCTACAACGCACGCCAGAGGGTTGCCCGCAGCCGCGCC 121

Qy 291 ThrProLeuLeuAspLeuIleMetAlaAlaLeuThrProHisProHisProGlnLysGlnAla 310
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122 ACCCCGTTATTAGATTTGATCAAGACAGCGGTTGAGCGCCCCATCCACCGCAAAAACAGGCG 181

Qy 311 TyrGlyValThrLeuProThrSerValLeuPheIleAlaGlyHisAspThrAsnLeuAla 330
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182 TATGGTGTGACATTACCCACTTCAGTGCTGTTTATCGGCGGACACGGATACTAATCTGGCA 241

Qy 331 AsnLeuGlyGlyAla 335
        ||||||||||||||
Db 242 AATCTCCGGCGGCGCA 256

Search completed: June 14, 2006, 15:26:30
Job time : 4992.33 secs

THIS PAGE LEFT BLANK

```
; FILE REFERENCE: MOA-A020SY1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 5791
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 27838
; LENGTH: 3731
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
;   DATABASE ACCESSION NUMBER: AK103280
;   DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-27838

Alignment Scores:
Pred. No.:
Score:                    18.6         Length:        3731
Percent Similarity:       95.00        Matches:         80
Best Local Similarity:    34.6%        Conservative:    52
Query Match:              20.9%        Mismatches:     150
DB:                        4.1%        Indels:         100
                           6           Gaps:            16

US-10-601-319-2 (1-440) x US-10-449-902-27838 (1-3731)

Qy     120 GluAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThr------- 136
               :::::|||         :|  |:::|||          ::
Db    1999 GATGCTTCCTTACTTGGTCTCTGCAACCTGACACATGTATCAAACGATAGCTGCGCTTTC 2058

Qy     136 ------------------------------------------------------------ 136

Db    2059 CAGGACCCAACAAAGCAGGTCTCTCCCACGATATTGCAGTTCCAGCCGCAGAACATA 2118

Qy     137 ----GlnAlaAspThrSerSerProAspProLeuPheAsnProLeuLysThrGlyVal 154
               :::|:||:::|:||||::|
Db    2119 GGTGGCAGAGCTAATACACTTCTGCCAAGTCAGATTTTGCAGCAAGTCCAGCCTCAGTT 2178

Qy     155 CysGln--------LeuAspAsnAlaAsnValThrAspAlaIleLeuSerArgAlaGly 171
               |||              :::
Db    2179 CAGCAGCAGCAATACCTCCAAAACATCAACGAGACC 2214

Qy     172 GlySerIleAlaAspPheThrGlyHisArgGlnThrAlaPhe------ArgGluLeuGlu 189
               |||        :||||:||:||  |||          :|||:|||:::
Db    2215 ---ACAATCCAAGGGACTGCTCAGTCTGAGTTCCTTCAGCAGCTCCAA 2262

Qy     190 ArgValLeuAsnPheProGlnSerAsnLeuCysLeuLysArgGlyLysGlnAspGluSer 209
               |||:||    :::|||||||||            :::||||:||
Db    2263 CGCTGCCAGTCGTTTACTGAGAGAAGCCACAGCTGCAAACCCAACAACAGCAAGAA 2322

Qy     210 CysSerLeuThrGlnAlaLeuProSerGluLeu------ 220
               |||
Db    2323 TCACAGCAAACAACAGCAGCGCAATCGCAGTGTATGCCACAACACATCAGCAAATG 2382

Qy     221 ------LysValSerAlaAspAsnValSerLeuThrGlyValAlaValSerLeuAlaSer 237
               :|          |:||
Db    2383 CAACAGCAGAAGAACAATGACCAACTATCAGTCTGTACCTAAGCAATTGTCACCGTTTTCT 2442

Qy     238 MetLeuThrGlu--------------------IlePheLeuLeuGln 246
               |||:|||       |||:::
Db    2443 CAAACTCTCCTCTCCCTTCTTCAGTCTTCACCTATGACACTGCAAACAGTATTACCATTCTCC 2502

Qy     247 GlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThrAspSerHisGlnTrpAsn 266
               |||        |||:::
Db    2503 CAGCCACAGAGCTATCCAGACACAGAGTATGAGCTCATTATACCATCC------AAC 2553

Qy     267 ThrLeuLeuSerLeuHisAsnLeuGlnPheTyrLeuLeuGlnAsnArgThrPro------ 283
               :::::|||::|||||||
Db    2554 ACA---TCCACCATGCATAATGCATTGAGGCCATTTCATCGGAAGCACCCTCGCACCTA 2610

Qy     284 GluValAlaArgSerArgAlaThrProLeuLeuAsp------------------LeuIle 297
               :::
```

Qy   298 MetAlaAlaLeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThr 317
               |||:::
Db   2611 AGCAATGCCAAGACCGACACTGCAGTACCTGTCCCCTGACCCATTGCGAGACGGAGTTGCA 2670

Qy   318 SerValLeuPheIleAlaGlyHisAspThrAsnLeuGluAlaAsnLeuGlyGlyAlaLeuGlu 337
               :::   ::::::
Db   2671 GTGGAGTCTTTGCTTCCCTCGTCCTCAGGTTACATGCAGGATGGAGCAGTTGGACTCA 2730

Qy   338 LeuAsnTrpThrLeuProGlnThrProGlyLysGluLeuValPhe 357
               ||| :::
Db   2731 ACAGCACCTAGTATATCCA---CAAAGCTCTGCGGTTGGGCGGGCTTTCCCTGGA------ 2778

Qy   358 GluArgTrpArgLeuSerArgAspAsnSerGlnThrProIleGlnValSerLeuValPheGln 377
               :::::| ::::::
Db   2779 ---AGAGGATGCTTGGCTGTGGATCAAGATGGGAAGCTCAAAACCATCGTTATTT------ 2835

Qy   378 ThrLeuGlnGlnMetArgAspLysThrProLeu----------------------- 388
               :::||||||| ||| :::
Db   2884 AGCCTTCCAGGGTGAGGAAATGAATTCGACGGCAGAATTCCTTATTCCACCTCCAGTTTCCTGAGC 2943

Qy   389 ---SerLeuAsnThrProProGlyGluValLysGlyLeuThrLeuAlaGlyTyrCysGluGlyArg 407
               ||||||| ::::::   :::||||:||
Db   2944 CCGTCGCAGAATGATTTCCCTTGGATCAAACACTAAGTAGTGCAGACTGCCTTAGATGAA 3003

Qy   408 AsnAlaGlnGlyMetCysSerLeuValAlaGlyPheThrGlnIleLeuValAsnGluAlaArgIle 427
               :::|||||| :::||||:||
Db   3004 TCAGGATATGTGCCGTGTTCA---CAGAATTCGATCAAGTGATTAAT-----CGACCA 3054

Qy   428 ProAla 429
               ||||||
Db   3055 CCTGCG 3060

Search completed: June 14, 2006, 14:24:53
Job time : 59.3073 secs
```

THIS PAGE LEFT BLANK

```
Db   1864 GCTGAACAAGAGAACTGGGGATTGAGGTTCAGACCCC------GTGCCCCCCGTGAAGTTGCA 1917
Qy    352 -GlyGlyGluLeuValPheGluArgTrpArgLeuSerAspAsnSerGlnTrpIleGl 371
Db   1918 GGGCGGCAATCCTCGCGTTGGTTTTCCTGGAATCACTGCCTCCCGTGGGTTGACTATGGTGAC 1977
Qy    371 nValSerLeuValPheGlnThrLeuGlnIleMetArgAspLysThr 386
Db   1978 AATCATCATGACGAAAATGACCAAAACAATGATGACTGATTCAACC 2023
```

RESULT 14
US-10-449-902-24588
; Sequence 24588, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24588
; LENGTH: 1488
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK100030
; DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-24588

Alignment Scores:
Pred. No.:
Score: 5.04
Percent Similarity: 95.00   Length: 1488
Best Local Similarity: 33.9%   Matches: 92
Query Match: 21.2%   Conservative: 55
DB: .4.1%   Mismatches: 133
                                     Indels: 154
                                     Gaps: 6

US-10-601-319-2 (1-440) x US-10-449-902-24588 (1-1488)

```
Qy      9 LeuSerLeuLeuIleProLeuThrProGlnSerAlaPheAlaGlnSerGluPro------ 26
Db     30 CTCGATCCCCTCCCCCTGTCTCCCCGATCCCCTTTCTCTCCCCTTCCCCTCCCGGCGGCCG 89
Qy     27 ------------------------GluLeuLysLeuGluSerValValIleVal 36
Db     90 CTTTCCTCGCCGCCGTCCCCTTCCGGGTGCGGCCGCTCCCCCCTTCCCCTACATCCCGACCTCC 149
Qy     37 SerArgHisGlyValArgAlaThrLysAlaProThrLysAlaThrGlnLeuMetGlnAspVal 54
Db    150 ACCTCCACCGCCCGCCGCCCTGCCGCCTCCTCCGCGAGGAGCACAGACCCTAGCGACCTTCCAATT 209
Qy     55 -------ThrProAspAlaTrpProThrTrpProValLysLeuGlyTrpLeuThrPro 71
Db    210 CGTCGCCGCCGCCGACTCGCCGGTCTCCACGCGTAAGAGCTCCCCCCC 269
Qy     72 ArgGlyGlyGlu--LeuIleAlaTyrLeuGlyHisTyrGlnArgGlnLeuVal 89
Db    270 TTCTCCCGGCGAGACGTGTCTTCTATGATTCTCCCCCATTGT--CGTCGCCGATTGGCC 326
Qy     90 AlaAspGlyLeuLeuAla------ 95
Db    327 GGCGATGGATTACTGGCTCCGTCTTCTTCCCGGGGGCGGGGACAACATCTTCGACGCCAT 386
Qy     96 ---LysLysGlyCys------ProGlnSerGlyGlnVal------ 105
```

```
Db    387 CGCCGTCGCGGCGTCCGACCACCCCGCCGCCGGCTCAGTGCAGCGGGACGCCATCGCCCA 446
Qy    106 AlaIleIleAlaAspValAspGluArgThrArgLysThrArgLysThrGlnAlaPheAlaAlaGly 125
Db    447 GCGCTTCTACACGGCGTACCGTTGGCGGTGCCTCCCAGCCGCCACCGCCGAGGCA 506
Qy    126 LeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThrSerProAspPro 145
Db    507 CCGCCGCCGAG------CTGCTCCACCCCGAGGGCGGCCGCTAGCGTCCCT----- 551
Qy    146 LeuPheAsnProLeuLysThrGlyValCysGlnLeuValAspAsnAlaThrAspAla 165
Db    552 ------------AGCATCGTAGCTCCGACCGCCAATGTC----- 581
Qy    166 IleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGly-------HisArgGln 182
Db    582 ------ATCGGCGACGACGGCGGCGTCCTTGCCATGAGGAC 617
Qy    183 ThrAlaPheArgGluLeuGlyValLeuAsnPheProGlnSerAsnLeuCysLeuLys 202
Db    618 CCCGTCCGCCGCGAGACCGAGCGCATC-----------AAGGCCGTCCTCCTC 659
Qy    203 ArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuPro------SerGluLeu 220
Db    660 AACGACCAAGAAAAGTCGGACGCGACGTTGCTGCAGCTGCTCCGGGTCTGCAGCAGCTG 719
Qy    221 LysValSerAlaAlaAspAsnValSerLeuThr-------GlyAlaValSer------ 234
Db    720 GAGCTCACGGTGGACACCTTGACGGTTACTGAGATTGGAAAGCGGTGAGCAGCTACCGG 779
Qy    235 ------------------LeuAlaSerMetLeuThrGluIlePheLeuLeu 245
Db    780 AAGCACAACTCCAAGCAGCAGATTCGTCACCTTGTTCAATTGCTCATCGAA---------- 827
Qy    246 GlnGlnGlnAlaGlnGlyMetProGluProGlyTyrArgIleThrAspSerHisGlnTrp 265
Db    828 GGTTGGAAGCGCATACTCGAT-----GAGTGG 854
Qy    266 AsnThr------LeuLeuSerLeuHisAsnAlaGlnPheTyrLeuGlnArgThr 282
Db    855 ATGAGCAGCTGAGAGCGCCATTGTAGTGCATTGGATGGAATTGTTACTTTGTGCAAGTGC 914
Qy    283 ProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAlaLeuThr 302
Db    915 AGCCAA------AAGGACCAGCCATCTCCTTGCTTA 944
Qy    303 ProHisProProGlnLysGlnAlaTyrGlyValThrArgSerValLeuProPheIle 322
Db    945 ------GCACAACAGATTGCTTTT 962
Qy    323 AlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGlyLeuAsnTrpThrLeu 342
Db    963 CTTCCTAAGCATAGGTGGCCTGAA-GATGTAGCTTGG 997
Qy    343 ProGlyGlyProAspAsnThrProProGlyLysGlyLeuValPheGluLeuArgTrpArgArg 362
Db    998 ------CATCCAGATGTGAATTAATTTTTGCAATGTACAGT--- 1033
Qy    363 LeuSerAspAsnSerGlnTrpIleGluValSerLeuVal 375
Db   1034 ---GCTGATAACGGCGGAC--TCTCAAGTTTCACTGATG 1066
```

RESULT 15
US-10-449-902-27838
; Sequence 27838, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF

```
Db  1205  GTGATAGGCTCCTCTCAGAGTCGGATCTATCACATCTCACCTTCTTCCATGCTATCATCA  1264
Qy   377  lnThrLeuGlnGlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluv  397
                :: ::::::                            ||||||||||||
Db  1265  AGGAGACATTCCGTCCTACATCCATCCATCAACACCGGCTCCTCGCTG-----------  1305

Qy   397  alLysLeuThrLeuAlaGlyCysGluGluValAlaGlnAlaGlnAlaGlnGlyMetCysSerLeuAlaG  417
                                              |||   |||    ||:::|||
Db  1306  --------------------------CCACGCATGGCATCTGAGGAGTGTGAGATCGGCAG  1339

Qy   417  lyPheThrGlnIleValAsnGluAlaArgIlePro  428
                ||||::                ||||||||
Db  1340  GCTAC-------------------CGTATCCCC  1353


RESULT 13
US-10-449-902-26929
; Sequence 26929, Application US/10449902
; Publication No. US2006012350SA1
; GENERAL INFORMATION:
;  APPLICANT: National Institute of Agrobiological Sciences.
;  APPLICANT:  The Institute of Physical and Chemical Research.
;  APPLICANT:  Bio-oriented Technology Research Advancement Institution.
;  APPLICANT:  Foundation for Advancement of International Science.
;  TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
;  FILE REFERENCE: MOA-A025Y1-US
;  CURRENT APPLICATION NUMBER: US/10/449,902
;  CURRENT FILING DATE:   2003-05-29
;  PRIOR APPLICATION NUMBER: JP 2002-203269
;  PRIOR FILING DATE: 2002-05-30
;  PRIOR APPLICATION NUMBER: JP 2002-383870
;  PRIOR FILING DATE: 2002-12-11
;  NUMBER OF SEQ ID NOS: 56791
;  SOFTWARE: PatentIn Ver. 2.1
;  SEQ ID NO 26929
;  LENGTH: 2293
;  TYPE: DNA
;  ORGANISM: Oryza sativa
;  PUBLICATION INFORMATION:
;  DATABASE ACCESSION NUMBER: AK102371
;  DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-26929

Alignment Scores:
Pred. No.:           8.35      Length:          2293
Score:              95.50      Matches:           99
Percent Similarity: 32.6%      Conservative:      69
Best Local Similarity: 19.2%   Mismatches:       180
Query Match:         4.1%      Indels:           168
DB:                     6      Gaps:              26

US-10-601-319-2 (1-440) x US-10-449-902-26929 (1-2293)

Qy     7  ProPheLeuSerLeuLeuIleProLeuThrProGlnSerAlaPheAlaGlnSerGluPro  26
              |||    :::::: |||:::     ||||:::                 ||||:::
Db   574  CCTGGGACCATCATCATTCTCTCAATTGATCCCAAACATCTAAAAGCTGAAAATGAGATG  633

Qy    27  GluLeuLysLeuGluSerValValIleValSerValIleValSerArgHisGlyValArg--AlaProThr  45
              |||||||    :::   ||||||    |||           ||||:::
Db   634  AGGCGCATTTTTGGATCAAAAGTAGTGGAATTCACTTGAAAATCAACGGAATGTGCCTAGC  693

Qy    46  LysAlaThrGlnLeuMetGlnAsp-----------------------------------  53
              :::||||||        :::::::
Db   694  AAACAGGATAACAGAGAGGCCCAGGACTCGGGCATGTGTTAGAAAAGAGCATTCTTCCTT  753

Qy    54  ---ValThrLeuProAspAla---TrpProThrTrpProValLysLeuGlyTrpLeuThrPro  71
                 |||||||  ||| |||         |||||:::
Db   754  CTTGTTACTCCATCAGCTACTGGCCACCATGGGATAAATCAATGTCAATG---------  804

Qy    72  ArgGlyGlyGluGluValIleAlaTyrLeuGlyHisTyrGlnArgLeuValAlaAsp  91
              :::   |||||||||    :::
Db   805  ---------GATATTGTGGAGACAAAGAGTGGTTCAACTATTTCAGGTACATATATGAT  855

Qy    92  GlyLeuLeuAlaLysLysGlyCysProGlnSerGlyGlyGlnValAlaIleAlaIleIleAlaAspVal  111
              ||||||
```

```
Db   856  ---------------------------------------------------CCTTCTGTCAGTCATGTGCAAGACTTG  882
                                                                   ||||:::: :::
Qy   112  AspGluArgThrArgLysThrGlyGlu-------------AlaPheLeuAlaAlaGlyLeuAla-----  127
              ||||                                     ||||||||
Db   883  TTTGAAGTTGCCAAATCTGCAAATGATCTGAGTGCTATTGCAGCTATATTAGCAAAGTAC  942
                :::||||:::
Qy   128  ----------ProAspCysAlaIleThrVal-------------HisThrGlnAlaAsp  139
                       |||||||||                       ::::::
Db   943  CCTTATCACCCAGAATCATTGTTAACATTTGCTGAACTTTTCAAATATTCTGAGAACAT  1002

Qy   140  ThrSerSerProAsp-------------------------------Pro  145
              ||||||||
Db  1003  CAATCGTCAGCAGATGCTGTTGAGAGTGTCTATTTGCTCTGGAGTGTGCTTGGCATCCA  1062

Qy   146  LeuPheAsnProLeuLysThrGlyThrGlyValCysGlnLeuAspAsnAlaAsnValThrAsp-----  164
              ||||:::|||||||                           :::: :::
Db  1063  TTGTTAGCCCACTGCACAGCAAC--TGCCAGTGAAATATAGCCATGACAACACAAACAAG  1119

Qy   165  -----------------------AlaIleLeuSerArgAlaGly------------  171
                                     |||   |||  |||
Db  1120  CCATTCTTTACTGCGGCTTTTAGTCACATGAAAAATTTGGATAGACGTGGCTGCCATCGG  1179

Qy   172  ------------------------------------GlySer  173
                                                  ::|||
Db  1180  TCTGCTTTAGAGGTCTGCAAATTCCTTCTTCATTGGATTCTGATGATCCAAAGGGTGCT  1239

Qy   174  Ile------------AlaAspPheThrGlyHisArgGlnThrAlaPheArgGluLeuGluArgVal  191
              :::                  |||:::                ::::::
Db  1240  CTATTCTGCATTGATTACTTGCTCTAAGATCACAACAGTACAAATGGTTGGACACAGTTT  1299

Qy   192  LeuAsn--------------------------PheProGlnSerAsnLeuCysLeu  201
              :::                            ||||||
Db  1300  GCAGAAGAGTACCAGTGTGATAACTCCTGTGGTTATTCCCTAATTCTCATTTTCCCTT  1359

Qy   202  --------------------------------LysArgGluLysGln  206
                                               :::
Db  1360  GCCATTGCACGGTTCACGTTGAGCGTGATGCAACAAATGGGAGGCGTTCTGACCATTCT  1419

Qy   207  AspGluSerCysSer--------------LeuThrGlnAlaLeu--------ProSerGluLeu  220
              ||||:::: :::             :::                    ||||||
Db  1420  GACAAGTCAACATCTAGATCTCATGAAGCAAGCATTGATGCTTCATCCTTTGGTGCTG  1479

Qy   221  LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr  240
              :::                  |||      :::
Db  1480  CGCAAGATTGTTGACAAGGCTCCTTGAAAGACTCC---TCGTGGACCCAAATACTCAGG  1536

Qy   241  GlyIlePheLeuLeuLeuGlnGlnAlaGlnAsnGlyMetProGluProGlyTrpGlyTyrArgIleThr  260
              ||||||||                   ::: |||
Db  1537  AATGTGTTCTTT-----------------------GGATCAGCAAAGCCAGGA-------  1566

Qy   261  AspSerHisGlnTrpAsnThrLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln  280
              :::::::::::::
Db  1567  ---AGCCCTTCCCTGAGCATATGATTAGCATATATGTTGAACGCCATTACATCATGTGG  1623

Qy   281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIle------Met  298
              ||||||  :::
Db  1624  AGGTTCCCAGAACTGCAGAATTACTCAAAGATGCTGCCTTTTGGTGATTGAATCACTA  1683

Qy   299  Ala-AlaLeuThrProHisPro----------------ProGlnLysGlnAlaTyrGl  312
              ||| |||                             |||
Db  1684  AAACAGGATAACAGAGAGGCCCAGGACTCGGGCATGTGTTAGAAAAGAGCATTCTTCCTTCA  1743

Qy   312  yValThrLeuProThrSerValLeuPheIle-----------------AlaGlyHisAspTh  327
              :::  :::                         |||
Db  1744  GATAAGAATGAGTACTCTCATCTTCTTGTTCAGACTTCTCCGACACACCATCACTC  1803

Qy   327  rAsnLeuAlaAsnLeuGlyGlyAla----------------------------Met  335
              ::: |||  :::
Db  1804  CCCCCTGAAGAATTGCGGCCATTCATCGTTGCTCCTGGAATGGCACATGAGATGCGCCA  1863

Qy   336  -LeuGluLeuAsnTrpLeuProGlyGlnProAspAsnThrProPro---------  351
              ||| |||
Db  1864
```

```
Db   268 GTG------------------------------CTGGGCAACCTGCCGCAGTCTGGCGGGAAGACGCACCAGACGCTG  315
Qy    75 ---GluLeuIleAlaTyrLeuGlyHisTyrGlnArgLeuValAlaAspGlyLeu  93
                 ||:::            |||:|||:||
Db   316 CACGGAGATGACCAAGGTACGGCCCCGATCCGGTCCGGTTCGGGAGCTCCGACGTG  375
Qy    94 LeuAlaLysLysProGlnSerGlyLnValAlaIleIleAlaAlaAspValAspGlu  113
                 :::            |||            :::
Db   376 GTGGTCGCCGGTCGGCGCCGGTGGCGGCGAGTTCTCCGCCACGATGCCAAC----  432
Qy   114 ArgThrArgLysThrGlyGluAlaPhe-AlaAlaGlyLeuAlaProAspCysAlaIleThr  133
                 |||||||            ::::|||||||
Db   433 ----------TTCAGCAGCCGGCCACCGGCCAACTCCGGCGGCGAGCAC  468
Qy   133 rValHisThrGlnAlaAspThrSerProAspProLeuPheAsnProLeuLysThrGl  153
                 |||            |||:::
Db   469 ATGGGCTACAACGCCCGGACGTCGTGTC-GGGCCG---TACGGCGGGTGGCGCGC  524
Qy   153 yValCysGlnLeuAspAsnAlaAsnValThrArgAspAlaIleLeuSerArgAlaGlyGlySe  173
                 :::     :::::::|||            |||:::
Db   525 CATGCGGAAGATTTGCCGCGTCAACCTCTTCCGCG---------CGCGGCCTCGACGA  575
Qy   173 rIleAlaAspPheThrGlyHisArgGlnThrAlaPheGluArgValLeuAs  193
                 :::                |||
Db   576 CCTG--------------------------CGCGCTTTCCGGAGCGCGGAGGGCCGTGCTG--  609
Qy   193 nPhePheProGlnSerAsnLeuCysLeuLysArgGluLysGlnAspGluSerCysSerLeuTh  213
                 :::            :::|||            |||||||
Db   610 --------ATGGTTAGGTCGCTGCGGAGGCGAGCGGCGCCCCCTGG  647
Qy   213 rGlnAlaLeuProSerGluLeuLysValSerAlaAsp---AsnValSerLeuThrGlyVal  232
                 :::      |||    :::          |||||||
Db   648 GTCGTCGTCTCCAGCGGCGGTGTCCTGGAAAGGAGTGAATGTCTGCACGACGAACAGC  707
Qy   232 aValSerLeuAlaSerMetLeuThrGlulePheLeuLeuGlnGlnAlaGlnGly------  250
                 ||||||            |||:::::::
Db   708 GCTGTCGCGCGCGCCGGTGTCGGGCGCCGTGTTCGCCGCCGGCGGCGGCGAGGGCGCGGAG  767
Qy   250 --------------------------------------------GlyArgIleThrArgSe  262
                                                            |||:|||
Db   768 GGAGTTCAAAGAGATCGTCGTCGAGGTGATGAGGTGGGTGGTGTGTGAACCTCGGCGCA  827
Qy   262 ---------MetProGluProGlyTrp----------------  262
                        :::::|||:::
Db   828 CTTCGTGCCGGCGTCCGGTGCGTGACCCGCAGGGCGTGGTAGCGAGGATGAAAAAGCT  887
Qy   262 rHisGln-----------------TrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTy  277
                 |||:::                            |||
Db   888 GCACCGCCGGTTCGACGACATGATGAACGGCGATCATCGCGAGAGAGGGCCGGATCACT  947
Qy   277 rLeuLeuGlnArgThrProGluValAlaSerArgAlaThrProLeuLeuAspLeuil  297
                 |||||||            ::::|||            |||||||
Db   948 ACTC-----------AAACCAACCGACAGTCGTGAGGAAGGTAAGGGACTTGCTTGGCTTGCT  998
Qy   297 eMetAlaAlaLeuThrProHisProProGlnLysGlnAlaTyrGlyValThr---------------  314
                 ::::|||            |||:::
Db   999 CCTGGCTATGGTG-----------------------------CAGGAGCAGGAGGTGGCTCCGCCCGCGAGGA  1043
Qy   315 ----------LeuProThrSerVal------------------------LeuPheIleAlaGlyHi  325
                 |||                   |||||||
Db  1044 CGACAGGATCACCGACCACGGAAATCAAGGCCCTTACCTCATTGCTCCTC  1103
Qy   325 sAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrpThrLeu----  342
                 :::            |||||||
Db  1104 TCAACCATAGTTGAGTGGACAAATGCAGAGCT  1145
Qy   343 ----------ProGlyGlnProAspAsnThrProProGlyGly-GluLeuValPheGluA  359
                 |||            ::::::|||
Db  1146 GATTCGACACCCAGA-TATCCTCAAGCACGCCCAAGAGAGCTAGATGTTGTTGTGGGTC  1204
Qy   359 rgTrpArgArgLeu------SerAspAsnSerGlnTrpIleGlnValSerLeuValPheG  377
                 |||            |||    :
```

```
Db   729 GGCGGTGTCGGAGGGGCCACAGGTTCGGCGTGTCCATCGGAGAGAGCCAGATCGTAGGCATC  788
Qy   360 TrpArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeu  379
                 |||            ||:|||:|||:||
Db   789 AGGAGCAGAACAGAGCTCGAGTCGGAGTCGGCTGCAGGTG----------CTT  830
Qy   380 GlnGlnMetArgAspLys-----ThrProLeuSerLeuAsnThrProProGlyGluVal  397
                 :::::::            |||            |||||||
Db   831 GGCAAGCTCTACAAGAGACCCGGTGATCCGGCCGTGTTCCCTGTCCCG------CCC  884
Qy   398 LysLeuThrLeuValAlaGlyCysGlu----------GluArgAsnAlaGlnGlyly  411
                 :::::|||||||            |||:::|||||||
Db   885 ACGCAAGACATCCCGGCCACAGAGCGACTTTGCGGTCGGCGTGGCGACAGACAGGCTCAGGGC  944
Qy   412 MetCysSerLeuValAlaGlyPheThrGlnIleValAsnAlaLysArgIleProAlaCysSer  431
                 ||:::::::||||||            :::::::
Db   945 TCCGTCGTGTACGGGCGGCCCTTCGGC-----------AGCGAGGGCGAAGCTGACGAGCGCGCAG  995
Qy   432 LeuArgSer  434
                 ||:::::::
Db   996 CTCCAAACA  1004
```

Alignment Scores:
```
Pred. No.:                 6.21        Length:              2010
Score:                    96.00        Matches:              112
Percent Similarity:       34.3%        Conservative:          57
Best Local Similarity:    22.7%        Mismatches:           174
Query Match:               4.2%        Indels:               152
DB:                           6        Gaps:                  24
```

US-10-601-319-2 (1-440) x US-10-449-902-20112 (1-2010)

```
Qy    11 LeuLeuIleProLeuThrProGlnSerAlaPheAlaGlnSerGluProGluLysLeu  30
                 ||:::::||:::            :::            |||
Db   106 CTAGTCCTACCA---ACAATGGAGGTCGCCCATGGAGATCTCTACCTCATTGCTCCTC  162
Qy    31 GluSerVal----------------ValIleValSerArgHisGly----  40
                 :::||||            :::::::|||
Db   163 ACACCGTGGCTCTCCGTCATCGTGTCCGTACGCCCTGGTCTTCCCGGCCGGGAAG  222
Qy    41 ValArgAlaProThrLysAlaThrArgGlnLeuMetGlnAspValThrProAspAlaTrpPro  60
                 |||            ||||:::||:|||:||
Db   223 GGCGGTGCGCG----------------CTGCCGCTGCGCCCTGGCCCGCCAGGGGATGGCCG  267
Qy    61 ThrTrpProValLysGlyLeuTyrTrpLeuThrProArgGlyGly----------  74
                 ||||||            |||    :
```

```
; Sequence 34901, Application US/10953349
; Publication No. US2006010734SA1
; GENERAL INFORMATION:
;   APPLICANT: ALEXANDROV, Nickolai et al.
;   TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
;   TITLE OF INVENTION: ENCONDED THERBY
;   FILE REFERENCE: 2750-1579PUS2
;   CURRENT APPLICATION NUMBER: US/10/953,349
;   CURRENT FILING DATE: 2004-09-30
;   NUMBER OF SEQ ID NOS: 40252
;   SOFTWARE: PatentIn version 3.3
;
; SEQ ID NO 34901
;   LENGTH: 1619
;   TYPE: DNA
;   ORGANISM: Zea mays subsp. mays
US-10-953-349-34901

Alignment Scores:
Pred. No.:              3.68          Length:        1619
Score:                 97.00          Matches:         81
Percent Similarity:    36.5%          Conservative:    37
Best Local Similarity: 25.1%          Mismatches:     107
Query Match:            4.2%          Indels:          98
DB:                        6          Gaps:            23

US-10-601-319-2 (1-440) x US-10-953-349-34901 (1-1619)
```

```
Qy      170  AlaGlyGlySerIleAlaAspPheThr-----------------------------GlyHis  180
                  ||||||                                              ||||||
Db      156  GCGGTCGGCACATCCTCCCCTTCACGGAGCTCGGCCAAGGCATAGCTAGGCAGGGCCAC  215

Qy      181  ArgGlnThrAlaPhe--------------ArgGluLeuArgValLeuAsnPheProGlnSer  197
                  |||||||||                 |||:::::||||||||||||||||
Db      216  CGGGTCACCCTCCTTCTCCACGCCAAGAACACCCGGCGGTAATCCGACCATCCTCCAGAA  275

Qy      198  ----------------------------AsnLeuCysLeuLysArg---GluLysGlnAspGlu  208
                                             ::|:::||||          |||:::
Db      276  CTCGCGGGGCATATCCGGTACCGGTCGTGGACATCGGCCGCGGGCCTGCCGGAG  335

Qy      209  SerCysSerLeuThrGlnAlaLeuProSerGlu----------LeuLysValSer  223
                  :::||                            |||||||:::
Db      336  GACTGGAGGGCGGAGCATAGACCTGTCCATCCAGACCTACCTCCGGTAGCC  395

Qy      224  AlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThrIleLeuPhe  243
                                                ::: |||||||||||||||||||----
Db      396  TACGAC--------GCCGCCTTCGCCGACAAACTCTCGGCCATC---  431

Qy      244  LeuLeuGlnAlaGlnGlyMetProGlu--ProGlyTrpGlyArgIleThrAspSer  262
                  ||| :::                         :::
Db      432  --CTGCAGGAGGCCGCGG-----CCGGAGGAGGCCGGACCGGGTTCTCATCGACTACGCC  482

Qy      263  HisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGlnArgThr  282
                  |||
Db      483  GCGTACTCG-----------------------------GCG  494

Qy      283  ProGluAlaAspSerGlnTrpIleGlnVal----LeuLeuAspLeuIleMetAlaAla  300
                  ::: |||                         ||||||
Db      495  CCCGGGCCGCCGGAGGACGCACGGCCGCTGCCGTCGGTTCGGGCCGCGCG  554

Qy      301  LysAlaTyrGlyValThr  314
                             :::
Db      555  ACCCTCAGTTCTCATGGGCCTCATGGGACGGGGCCTCATGGGACGGGGCCTCATGGGACAGGACG  614

Qy      315  AspThrAsnLeuAlaAsn  331
                  :::|||
Db      615  AAGCCGGAGGACCTCACCGTAGTACCCGACTACGTGCCCAACAACCGTCGCAC  674

Qy      332  Gly--GlyAlaLeuGluLeuAsnTrpThrLeuProGly--GlnProAspAsnThr  349
                  |||
Db      675  CGGCGGCCTTCGAGGCGCGCGGAGCTA---------TTCAAGCCAGGGCCTTGTCCCGGACTCC  728

Qy      350  ----------------------------ProProGlyGlyGluValPheGluArg  359
                                                     :::
```

```
Db      407  CGCCTCAGCGTCAAGCAGCAAATAGGCGCTACGGCCAGTTGACTCTG----------GTCAAC  457

Qy      113  GluArgThrArgLysThrGlyIleGluAlaPhe-AlaAlaGlyLeuAlaProAspCysAlaIl  132
                  :::           ||||||||                           |||||||||
Db      458  TCCACCTCGGCACAGACACAGGTGAA---TTCAGCTGCTGGG-----------TGCAGCTC  502

Qy      132  GluThrValHisThrGlnAlaAspThrSerPro-AspProLeuPheAsnProLeuLysT  152
                              ::: |||                     :::
Db      503  TGCAGCGGCTACATCTGCAGGAGGACGAGGCCCAAAACGGGCTC----------  546

Qy      152  hrGlyValCysGlnLeuAspAspAlaAsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlyG  172
Db      546  ----------------------------------------------  546

Qy      172  lySerIleAlaAspPheThrGlyHisArgGlnThrAlaPheArgGluLeuGluArgValL  192
                                           ::: |||||||
Db      547  ------------CACCTTACATCTTTTTACAGGAAGGAGAACTCTTT  583

Qy      192  GluAsnPheProGlnSerAsnLeuCysLeuLysArgGluLysGlnAspGluSerCysSerL  212
                  |||            |||||                |||
Db      584  GTACCTTCTCCCAGCTACTTCGATGTTGTCTACTTGAACCCGACAGACAGGCTGTGT-  642

Qy      212  euThrGlnAlaLeuProSerGluLeuLysValSerAlaAspAsnValSerLeuThrGlyA  232
                                                            :::|||||||
Db      643  ------------------------------------TCCTTGTCGGGTG  655

Qy      232  laValSerLeuAlaSerMetLeuThrIlePheLeuLeuGlnAlaGlnGlyMetP  252
                  ::::::                               |||||
Db      656  ACCGTGCTGTCGGCCCAAAGT--------CACGCGTCCACAGGGAATTC  694

Qy      252  roGluProGlyTrpGlyArgIleThrAspSerHisGlnTrpAsnThrLeuLeuSerLeuH  272
                  |||            |||||||
Db      695  CCAGCCAAGGA------------GATCCCCAGCCAATGGAACGGGACCATTGTTTATGAC  739

Qy      272  isAsnAla-GlnPheTyrLeuLeuGln  280
                  |||||:::
Db      740  ATGAAGCCGGGGCTTTGTGTATCTGCAACCTCATTCCGAGCACCCAGGGTGTGGTTTACTGC  799

Qy      281  ArgThrProGluAlaValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla  300
                  |||                      ||||||  :::
Db      800  AGGGCGGAGGACCGCGGGCAGATCCCAGATCTCCGTCAAGTACCAGCTGCTCTCTACGTGCG  859

Qy      301  LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu  320
                  :::                          :::|||
Db      860  GTTCCCAGTTGGATCTTCCCAGGGCAGAAGGATGAAAGGCT---------GTGACG  907

Qy      321  PheIleGlyHisAsp-----------ThrAsnLeuAlaAsnLeuGlyGlyAla  335
                  |||||:::                   ||
Db      908  GTGAAAAGTGGGGACGACCAGTGTGCTCGGGGGAGCCCGATGTGGAG  967

Qy      336  LeuGluLeuAsnTrpThrLeuProGlyGlnProProGlyGlyGluLeu  355
                  |||||||:::
Db      968  GTGGAGTTCACCTGGATCTTCCCAGGGCAGAAGGATGAAAGGCT---GTGACG  1018

Qy      356  ValPheGluGluTrpArg--------ArgLeuSerAspAsnSerGlnTrpIleGlnVal  372
                  ::: :::                  ::::::
Db      1019 ATCCAAGACACTTGGAGGTTGATCCAGACTGGGACACCACCAGAATCTCCCAG  1078

Qy      373  SerLeuVal---------PheGlnThrLeuGlnMetArgAspLysThrProLeu  388
                  ||:|:::         ||||||||
Db      1079 AGTGTCATTACAGTGGAAGACTTCGAGACGATTGAT------------------  1114

Qy      389  SerLeuAsnThrProProGlyGlyGluValLysLeuThrLeuAlaGly  405
                  |||||:::             |||||
Db      1115 AAGCCGGAGGACCTCACCGTAGTACCCGACTACGTGCCCAACAACCGTCGCAC  1135

Qy      406  GluArgAsnAlaGlnGlyMetCysSerLeuAla  416
                  :::|||                ::::::
Db      1136 GCTCAGAATCTTCAAGGACAGACCAGTAGCT  1168
```

RESULT 11
US-10-953-349-34901

```
; SEQ ID NO 570
;   LENGTH: 14756
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-505-928-570

Alignment Scores:
Pred. No.:                      61.3      Length:          14756
Score:                          98.50     Matches:            87
Percent Similarity:             34.9%     Conservative:       50
Best Local Similarity:          22.1%     Mismatches:        140
Query Match:                     4.3%     Indels:            116
DB:                                       Gaps:               17

US-10-601-319-2 (1-440) x US-10-505-928-570 (1-14756)

Qy     102 SerGlyGlnValAlaIleIleAlaGluAspValAspArgThrArgLysThrGlyGluAla 121
               |||          :: ::  :||         :: ::||||
Db   12432 TCCTCACACTGGCAAACAAAGCCCCGTTGTCGACAACACACGTGCCCCATAGAGGCA 12373

Qy     122 PheAlaAlaGlyLeuAlaProAspCysAlaAlaThrValHisThrGlnAlaAspThrSer 141
               |||               :::||||||
Db   12372 TGGGGTTGGAGGAACACGGATTGACGGTTATCTCACAGTGGGTCCCTATGTACAAGGCACT 12313

Qy     142 SerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAlaAsn 161
               :::    ::  ::           :: :::     ||||||
Db   12312 GCA---------TTTGCAGTAATAACCTCCAGCAGGTGACGGATTGCAAACGCC-- 12268

Qy     162 ValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHisArg 181
               ::::||||||        :::
Db   12267 -------------TCCATTCTGGCAAGGGTGCTGCGCAGTC------TTCCGTGGCCGTCAG 12223

Qy     182 GlnThrAlaPheArgGluLeuGluArgValLeuAsnPhe---ProGlnSerAsnLeuCys 200
               :::||||:  ::
Db   12222 GAAGCAGCCTGAGATACATCCACCGACTCTTCGATGTGTGCATAGCTTCTGGGGTTGCT 12163

Qy     201 LeuLysArgGluLys---------GlnAspGluSerCysSerLeuThrGlnAlaLeuPro 217
               ::::|||||                     |||
Db   12162 GTTTAAAGGGAGGCTCCTGCCCATTCAAATAAATGGAGTCCATACAACCCCTGAAACCATT 12103

Qy     218 SerGluLeuLysValSerAlaAspAsnAlaValSerLeuThrGlyAlaValSerLeuAlaSer 237
               ::||||:  ::
Db   12102 ACCAACTTGAGGACTTCTTCCATGCCTTGTTCCCTG-----------CTGACGGAT 12058

Qy     238 MetLeuThrGlnIlePheLeuLeuGlnAlaGlnGlnGlyMetProGluProGlyTrpGly 257
               ::: |||:  ::|                      |||       |||
Db   12057 GTGGGCCACCAAAAACACATAGTTATCCAGGTTCAGGGTTTTCAGATCCC---TGGGGC 12001

Qy     258 ------------------------ArgIleThrAspSerHisGln 264
                                       ||||||
Db   12000 TGTGCCGATGCAGTGATGAACTTGGTCTAGAACCAAGCGAGCATAGTTCCATTCACTTC 11941

Qy     265 TrpAsnThrLeuSerLeuHisAsnAlaLeuPheTyrLeuLeuGlnArgThrProGlu 284
               :::
Db   11940 CAGGGCCACTGCGTGCACTGCCCATCATTGACCTGAATGCTCTGAACAGAGAC----- 11887

Qy     285 ValAlaArgSerArgAlaThrPro------ 292
               |||||||
Db   11886 -----AATTCCAGGGCCACTTCCACAGTCAAACTTGTACTGCAGCCTTCCATGATGAAT 11833

Qy     293 LeuLeuAspLeuIleMetAlaAlaLeuThrProHisProProGlnLysGlnAla-TyrGl 312
               |||  :::::::::  :::  ::
Db   11832 CTCCAAGATGCTATAGTCAGTTCCTCGAGCATACATGACAACCGACATGCGTGGAATATGT 11773

Qy     312 yValThrLeu--------------------------------- 315
               :::
Db   11772 TCTGAGCCTCATGGTCAGTTCATCCTCTAATTGTTTCATTTCCGTCAGACGGTATTT 11713

Qy     316 -------ProThrSerValLeu--PheIleAlaGlyHisAspThrAsnLeuVal 330
               :||||:   ::  :::
Db   11712 CACGTAGCTGTTTCCAGTCAGTGTCATAGAGATGAACTCCCTGGGCACTGACCAAACCTGCC 11653

Qy     330 aAsnLeuGlyGlyAlaLeuGluLeuAsnTrpThrLeuProGlyGlnProAspAsnThrPr 350
```

```
Db   11652 GCTG-----GGACAGACACAGGTGTGTTTCCTCCCCAGGATCAGACACACCATTCGGA 11599
                   ||         :::::
Qy     350 oProGlyGlyGluLeuValPheGluArgTrpArgArgLeuSerAspAsnSerGlnTrpIl 370
               :: :||         ||  ||
Db   11598 TCCCTCAGGG------CACGGATCATCTTCACAGCCATGGTG 11563

Qy     370 e-GlnValSerLeuValPheThrLeuGlnIleMetArgAspLysThrProLeuSerL 390
               ||||||             ||||||
Db   11562 GACAGGTGGGCACCTTC------CCTCTTTGCA 11536

Qy     390 euAsnThrProProGlyValLysLeuThrLeuAlaGlyLeuThrAlaCysGluArgAsnAlaG 410
               :::||||||
Db   11535 GAGACACACCG------CTGCCCGTGTGGTGGCGGGAGTCACAAAACTCAG 11491

Qy     410 lnGlyMetCysSerLeuAlaGlyPheThrGlnIleIleValAsnGluAlaArgIleProLeuC 430
               :::|||||
Db   11490 TCTGGCTGTGC------T 11479

Qy     430 ysSerLeuArgSerHisHisHis-HisHisHis 440
               :||||:
Db   11478 GTGTGTTGACATCACACTTTCATCCACAGACAC 11446
```

RESULT 10
US-11-297-134-55
; Sequence 55, Application US/11297134
; Publication No. US20060111297A1
; GENERAL INFORMATION:
;   APPLICANT: Genzyme Corporation
;   APPLICANT: Roberts, Bruce
;   TITLE OF INVENTION: BLOOD FACTOR DOMAINS
;   FILE REFERENCE: 5270C
;   CURRENT APPLICATION NUMBER: US/11/297,134
;   CURRENT FILING DATE: 2005-12-08
;   PRIOR APPLICATION NUMBER: PCT/US2005/018461
;   PRIOR FILING DATE: 2004-06-09
;   PRIOR APPLICATION NUMBER: US 60/477,291
;   PRIOR FILING DATE: 2003-06-09
;   NUMBER OF SEQ ID NOS: 64
;   SOFTWARE: Patentin version 3.3
SEQ ID NO 55
;   LENGTH: 1502
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-297-134-55

Alignment Scores:
Pred. No.:   3.3   Length: 1502
Score: 97.00   Matches: 90
Percent Similarity: 31.4%   Conservative: 46
Best Local Similarity: 20.8%   Mismatches: 158
Query Match: 4.2%   Indels: 139
DB: Gaps: 18

US-10-601-319-2 (1-440) x US-11-297-134-55 (1-1502)

Qy   13 IleProLeuThrProGlnSerArgAlaPheAlaGlnSerGluProGluLeuLysLeuGluSer 32
               ::   ::   :: :: |||
Db   206 ATTCCTAAAATGAAGACAGGGACTCAGCCAATTCAGCACCAAAGACGAGTCTATCATG 265

Qy   33 ValAlaIleValSerArgHisGlyValArgAlaProThrLysAlaThrGlnLeuMetGln 52
               ::: :::::: :::
Db   266 ATGCAAGTGCTGGATAAAGGTCGTTCCAGAAACCGCCGTCACCGTGAGTCTGTGCGCG 325

Qy   53 AspValThrProAspAlaTrpProThrTrpProValLysLeuGlyTrpLeuThrProArg 72
               :::
Db   326 GGGCAAACTGTAGAGCTT---CGATGTAAAGGAGTAGAAGTTGGGTGGAGCTACCCCT--- 379

Qy   73 GlyGlyGluLysIleLeuIleAlaThrTyrLeuGlyHisTyrGlnArgGlnArgLeuAlaAspGly 92
               :: ::: ::
Db   380 ---------------GCGTATCTCGGACACCTTTAAG-----------GATTCT 406

Qy   93 LeuLeuAlaLysLysGlyGlyCysProGlnSerGlyGlnValAlaIleIleAlaAspValAsp 112
```

```
;   APPLICANT:   Bio-oriented Technology Research Advancement Institution.
;   APPLICANT:   The Institute of Physical and Chemical Research.
;   APPLICANT:   Foundation for Advancement of International Science.
;   TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
;   FILE REFERENCE: MOA-A0205Y1-US
;   CURRENT APPLICATION NUMBER: US/10/449,902
;   CURRENT FILING DATE:  2003-05-29
;   PRIOR APPLICATION NUMBER: JP 2002-203269
;   PRIOR FILING DATE:  2002-05-30
;   PRIOR APPLICATION NUMBER: JP 2002-383870
;   PRIOR FILING DATE:  2002-12-11
;   NUMBER OF SEQ ID NOS: 56791
;   SOFTWARE: PatentIn Ver. 2.1
;   SEQ ID NO 12801
;   LENGTH: 3287
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   PUBLICATION INFORMATION:
;   DATABASE ACCESSION NUMBER: AK110149
;   DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-12801

Alignment Scores:
Pred. No.:             6.51         Length:          3287
Score:                99.00         Matches:          104
Percent Similarity:   32.7%         Conservative:      56
Best Local Similarity:21.2%         Mismatches:       174
Query Match:           4.3%         Indels:           156
DB:                       6         Gaps:              20

US-10-601-319-2 (1-440) x US-10-449-902-12801 (1-3287)
```

```
Qy     1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla    20
           :::|||:|| ::|||   ||  ||  |:|| :|||
Db   334 CTGCAAAGCGCACTTCCGCCTTCCATCTCCCAGGCGATGCCGGCCAATCCGACCTCA---   390

Qy    21 PheAlaGlnSerGluProGlyLeuLysLysGluSerValValIleValSerArgHisGly    40
            ||||    ||| | |||   |||||||| :|||      ::|||
Db   391 ------AAAGGGCCTGCTTCTGCCATCTTCAGTCAGACGACTCGATCAACA          441

Qy    41 Val--ArgAlaProThrLysAlaThrGlnLeuMetGlnAsnValThrProAspAlaTrpP    60
             ||  |||||||:|||:::::|| |||  ::::|||   ||:::|||
Db   442 GTCCTCGAGACCACCCCGTTCCATGACCACAGAGCCATCCCAGGAGCGGCAAC          501

Qy    61 roThrThrProValLysArgGlyTrpLeuGlyThrProArgGlyGlyGluLeuIleAlaTyrL    80
             |:||||   |||   ||| ||| |||||:||||
Db   502 CA------------CTACCCAACAACAGGAGCTGCAGGGCTCAAG          534

Qy    81 euGlyHisTyrGlnArgGln                                           88
           ::
Db   535 TCTCCGACCTCACCAACCAGGTGACCGGTCTCAACGGAGAAGCTCGTCAAATCCTTCAATC   594

Qy    89 euValAlaAsp                                                    91
            :::|||||
Db   595 GCATCGCCGCGACCTCGAGGACGACTACTCGGACGCCCAGAGACCAATTCGCCACCATGTCGG   654

Qy    92                                                          GlYL    93
                                                                    |||
Db   655 CAAAGATCAAGGAGCTCGAAAAGGAGCGCGCAGAGCACCTCGACGCCCTCAACACTGGCC   714

Qy    93 euLeuAlaLysGlyCysProGlnSerGlyGlnValAlaIleIleAlaAspValAspG    113
           ||||   ::|||         :::::|::
Db   715 TCCTCGTCGAAAAGGCTCACGTCAGCAGCGAGATGAGGAGGATGATCGACCGGTCATCG    774

Qy   113 luArgThrArgLysThrGlyThrGluAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleT    133
           ||| | ||| :::|||||||||
Db   775 AAGAGACCGCCCAACGGCCAGGCA----------------------------------    800

Qy   133 hrGlnHisThrGlnAlaAspThrSerSerProLeuPheAsnProLeuProLeuLysThrG    153
           :::||||||  ||:   ::
Db   801 ---------GAGAGTGACAAGACAAGATCGAGGCGCCGAGCTCGAGCTCGA    849

Qy   153 lyValCysGlnLeuAspAsnAlaAsnValThrAspAlaIleLeuSerArgArgAlaGlyCysLys    173
```

```
Db   850 GTCTCTTTAGCGAAGCCAACAACAAGATGGTCGGCCGGCCTGTCGCGTGCCAGAGCGCG    909
             :::
Qy   173 erIleuAlaAspPheThrGlyHisArgGlnThrArgAlaPheArgGluLeuGlnArgValLeuA    193
           |||||||||||   |||   ::: ||
Db   910 AA------GAAAGAGCCCAACATGGAGGAGGCGGCTCAAGATACCGAGGGCATCATG-    962

Qy   193 snPheProGlnSerAsnLeuCysLeuLysArgGluLysGlnAsnGluSerCysSerLeuT    213
           |||   ::: ||| |||
Db   963 --CTCGAGCAACAGAAGGTGCTTGCCGATCTGCAGCGTGAAAAG---------------C   1011

Qy   213 hrGlnAlaLeuProSerGluLeuLysValSerAlaAlaAspAsnValSerLeuThr------    230
           |||:   :::|||:::::|||||:
Db  1012 TGCAAGCCGATCCAGCCTCCAGAAGATCACCGCCGACACTGACCTGCATCCATCTCCCCG   1071

Qy   231 ----------------GlyAlaValSerLeuAlaSerMetLeuThrG    241
                           |||||||:::::|
Db  1072 AAAGCGACCCGTCAGCTTCCATCGCCTCAACACTCCTCCTCGCTCGATCCATCTCCC   1131

Qy   241 luIlePheLeuLeuGlnGlnAlaGlnGlnGlyMetProGlu----------ProGlyTrpProGlyA    258
           :: ||||||| ||: :::                         |||
Db  1132 AGAGGACCCGCTGTGCGTGACGTCAAGCACACAAACCCGCTGCCGACCTTCCAGCCCAAGCTC   1191

Qy   258 rg----------IleThrAspSerHisGlnThrThrLeu------------    268
               |||                  ::: :::
Db  1192 GTCTCATCATCAACATGTCCCCTACCACGAGTTCCGTTCCTTCATCGCCCATCCGCGCA   1251

Qy   269 --LeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGlnLeuArgThrProGluValAlaArgS    288
            ||||||          ||||||||||
Db  1252 AATTGCGAAAACAGCTGGCTCCTTCTACAACTATCCA---CTCCCGCACGCGGCTCGCCG   1308

Qy   288 erArgAlaThrProLeuLeuAspLeuProLeuIleMetAlaAlaLeuThrProHisPro--------    305
           |||                                                ||||||
Db  1309 ACTCGCTCACCGGC-------------------------ACCCCTACCCCACGCACCG   1341

Qy   306 --------------ProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeuProPheI    322
                         ||||| ||| ||:   |||: :::
Db  1342 CTTCGCCAGCGCCCTCAGCAACCGTCTAGCCGGTATCGGCGACGACCAACTCATCGCGA   1401

Qy   322 leAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaAlaGluLeuAsnTrp-----    340
           ||||||                 ::: |||||||   :::
Db  1402 CTGCTTCA---TCCGCCTACAGCAAGGCATTGGCGGCGGTCTCTATTCCAGCTCGCCAT   1458

Qy   341                              ThrLeuProGlyGlnProAspA    348
                                        |||||
Db  1459 TCGCCATCGCAGGGGTCAGCCGACCGCGACCATCCTACTTTTACCGTCCAACGTTGACA   1518

Qy   348 snThrProProGlyGlyLeuValPheGluArgTrpArgLeuSerAspAsnSerG    368
           |||                        :::
Db  1519 AC-----------------------------CTGGTA--------------TCGC   1530

Qy   368 lnTrpIleGlnValSerLeuValPheGlnInThrLeuGlnMetArgAspLysThrPro-    387
           ||||   |||:::||||||||||||
Db  1531 TGCCAAATCAATTGTCCTTGCCCTTCATCAAGCGTCCCAGGAAGAGGATGTAGAACCGT   1590

Qy   388 --LeuSerLeuAsnThrProProGly    395
             |||:::  |||||||
Db  1591 GCCTCCCGTCTCGACTTCGGCGCCGGGGG   1616
```

```
RESULT 9
US-10-505-928-570/c
;   Sequence 570, Application US/10505928
;   Publication No. US2006008532A1
;   GENERAL INFORMATION:
;   APPLICANT: Ludwig Institute for Cancer Research et al.
;   TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
;   FILE REFERENCE: 28967/39178
;   CURRENT APPLICATION NUMBER: US/10/505,928
;   CURRENT FILING DATE:  2004-08-27
;   PRIOR APPLICATION NUMBER: US 60/363,019
;   PRIOR FILING DATE:  2002-03-07
;   NUMBER OF SEQ ID NOS: 866
;   SOFTWARE: PatentIn 3.2
```

```
Qy    273  snAlaGlnPhe-----------------------------TyrLeuLeuGlnA  281
           :::::                                    |||
Db    713  CATCCACAAAAAGAGGCTAGTCCTCGTCTCCACCTCCACAGGCCGCACTCCC  654
           ::|||||  |||:|| ||||||  ||||||
Qy    281  rgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAlaL  301
           :||||||| ||||||| |||:
Db    653  AGACCCCTGGGGTGGGCCCTTGCCCGGCACCGCCGCCCCTTGCAGCCAGGGTCCAGGCGAAAG  594
           |||||| ||||
Qy    301  euThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeuP  321
           ||||||
Db    593  CAGAGACGT-CATCCA-----------GGGGTACTGCCCCACGGACGGACCCGGT  553
           :::
Qy    321  heIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrpT  341
           :::::
Db    552  ATCTGGAAACACCACTCCGACCTCTCTG------------------  527
Qy    341  hrLeuProGlyGlnProAspAsnThrProProGlyGluLeuValPheGluGlyArgTrpA  361
           ||||||
Db    526  ------------------CCCGTGGGAGGAGGACGGGCATTTCCTCGTGATT  493
           ||||
Qy    361  rgArg-LeuSerAspAsnSerGlnTrpIleGlnIleValSerLeuValPhe  376
           ||| |||
Db    492  CCAGGGTCGTGTGCCACATGTCATCTTGGTGTCAGGGAACCCCTCGAGTTC  445
```

RESULT 7
US-11-217-529-274
; Sequence 274, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
;   APPLICANT: SUNTORY LIMITED
;   APPLICANT: NAKAO, YOSHIHIRO
;   APPLICANT: NAKAMURA, NORIHISA
;   APPLICANT: KODAMA, YUKIKO
;   APPLICANT: FUJIMURA, TOMOKO
;   APPLICANT: ASHIKARI, TOSHIHIKO
;   TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;   FILE REFERENCE: S-38-285
;   CURRENT APPLICATION NUMBER: US/11/217,529
;   CURRENT FILING DATE: 2005-09-02
;   PRIOR APPLICATION NUMBER: US 10/932,182
;   PRIOR FILING DATE: 2004-09-02
;   NUMBER OF SEQ ID NOS: 197023
;   SOFTWARE: Patentin version 3.3
; SEQ ID NO 274
;   LENGTH: 2259
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-11-217-529-274

Alignment Scores:
Pred. No.:   2.47   Length: 2259
Score: 101.00   Matches: 93
Percent Similarity: 35.3%   Conservative: 45
Best Local Similarity: 23.8%   Mismatches: 138
Query Match: 4.4%   Indels: 116
DB: 7   Gaps: 21

US-10-601-319-2 (1-440) x US-11-217-529-274 (1-2259)

```
Qy    101  GlnSerGlyLeuValAlaAlaIleIleAlaAlaAspValAspGluArgThrArgLysThrGlyGlu  120
           ||||||  :::::||| |||     ||| :::
Db      7  CAGAGCCAACAAGATCGCCTACGTGCTGAACTCCGAC-----ACGGCAGGAGGCCGCGAG  60
Qy    121  AlaPheAlaAlaGlyLeuAlaProAsp---------CysAlaIleThrValHisThrGln  137
           |||||| |||| |||||| ||| :::: |||
Db     61  TCGGCCTCCGTCGCGCTGCTTCGACAACAGAGCAGCTGGCCAAGCTGCTCGTCCACACG---  117
Qy    138  AlaAspThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeu  157
           ||||||
Db    118  ------------------CTCAAG-----------GAATTG  129
Qy    158  AspAsnAlaAsnValThrAspAlaAsnAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPhe  177
```

```
Db    130  GGCTACGAGACTCCGCCGGACCAGTGCTCCTCGGAGAGCGGCGGG---TACCAGAACGAG  186
           |||         :::|||||
Qy    178  ThrGlyHisArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSer  197
           :::|| |||||| |||||||
Db    187  TCGAACCATATCCAGACCTTCTC----------AGGCTGTCAAGACCCGGCCAGTC  234
           ||| :::
Qy    198  AsnLeuCysLeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaAlaLeuPro  217
           :: |||  |||||| |||
Db    235  GACCTGTTCGCTGCCCTGCCCTCGCACATAGC  282
Qy    218  SerGluLeuLysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSer  237
           |||||||
Db    283  TCGCCCGCTCAGGCTGCCGACTCGTGCAAACGCTGCTCTCGCCCGCAGCGCCACCAGCAGC  342
Qy    238  MetLeuThrGluIlePhe------LeuLeuGlnGlnAlaGlnGly-------------Met  251
           ::|||||||||    :::||||||||
Db    343  GCCACCGGGTCGTGTTTGACCACATGCTTCAGCAGCTGCAGGGACTGCAGCTGCTGATG  402
Qy    252  ProGluProGlyTrpGlyArgIleThrAspSer-----------HisGlnTrpAsn  266
           ||| :::::::
Db    403  GGGTCGTGGGGGCCCGTCCACTTGCTCCGGACGGCCGAGGTCACCACGCTCAGGAACTACGTC  462
Qy    267  ThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGlnLeuArgThrProGluValAla  286
           ::::::      |||  |||||||
Db    463  GAGATCATGATCCTCGTCCAACAGGCAGATCTTCCTCGAGTTCTTCCACCCGCTCGAGCT  522
Qy    287  ArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAlaLeuThrProHis-------  304
           :::                                              |||||||||||
Db    523  GCTGCT----------------------------GCTGCCGCCGCCCCCCACGCGCGGC  552
Qy    305  ProProGlnLysGlnAlaTyrGlyValThrLeuProThr----------  317
           |||| ||||||
Db    553  CCGCCC------------------ACCAGCGTCGCCCGTCCTGTACCTGCCGCAAGGTCCTC  594
Qy    318  ------------SerValLeuPheIleAlaGly--------------HisAspAsp  326
           :::::::::
Db    595  AAGTGCTTCATCGAGATCTGGACGACGCGCTGTTGGTGTCCAATGACCAGTTCCTCAACGAG  654
Qy    327  ThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrpThrLeuProGlyGlnPro  346
           ||||||                  :::|||
Db    655  GAGAACATCTTCAACCC-GGAAACAACCTTGAGAGAGCTCTCCACGTACCTCACGAACCC  713
Qy    347  Asp---------------------------AsnThrProProGlyGluLeuValPhe  357
           :::
Db    714  GAAGCTCACGGCGACCACCTGATCGAGCGCGACCACCATCTCCAAATA  773
Qy    358  GluArgTrpArgArg-----------------LeuSerAspAsnSerGlnTrpIle  370
           ||||:::
Db    774  CATCGACCCGAACGAGCTGCGTCCCAAGGGCCGCCTCTTGAAGCAGGAGGCCAT  833
Qy    371  GlnValSerLeuValPheGlnThrLeuGlnMetArgAsp------LysThrProLeu  388
           ||||||         ||| :::
Db    834  CAAGTACCAGCAGTCCAGGACATCTTCAACATCTCCGACCCGGACGACGCCCCCTT  893
Qy    389  SerLeuAsnThrProProGlyGluValLysLeuThrLeuAlaGlyCysGluGluAsn  408
           ||||||                    :::
Db    894  CTCCTCCGAGTCCCACCGCATCAA-------------------TCTGCTGCGAGGA----------CAACATCTC  938
Qy    409  AlaGlnGlyMetCysSerLeuAlaGlyPheThrGlnIleValAlaAsnGluAlaArgIlePro  428
           ||||:::
Db    939  GCACGACTTGACTGTGACCTTCCAGGA----------------  965
Qy    429  AlaCysSerLeuArgSerHisHisHisHisHis  439
           ||||| :::
Db    966  -----CTGGAAGACCATCCAGGACACCAC  989
```

RESULT 8
US-10-449-902-12801
; Sequence 12801, Application US/10449902
; Publication No. US2006012350A1
; GENERAL INFORMATION:
;   APPLICANT: National Institute of Agrobiological Sciences.

```
;  NUMBER OF SEQ ID NOS: 5458
;  SOFTWARE: PatentIn Ver. 2.1
;  SEQ ID NO 1957
;  LENGTH: 1945
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-11-293-697-1957

Alignment Scores:
Pred. No.:                              Length:      1945
Score:                 1.79             Matches:       95
Percent Similarity:  101.50             Conservative:  43
Best Local Similarity: 34.7%            Mismatches:   148
Query Match:           23.9%            Indels:       114
DB:                     4.4%            Gaps:          13
                          7

US-10-601-319-2 (1-440)  x  US-11-293-697-1957 (1-1945)
```

Qy   178 ------------------------------ThrGlyHisArgGlnThrAlaPheArgGluLeu 188
Qy   5 LeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAlaPheAlaGlnSer 24
Db   4168 CATGGGGTCCACGGCTTAGCAGTGGGGACAGGCCCATGTCCAGGTGGCCCCACACGCCCTG 4227
Db   1372 CTGATTCCGTTT----------------------CTAGTGGTGCCAAGACGCCGCCGCTCAGCT 1331

Qy   189 Glu---ArgValLeuAsnPheProGlnSerAsn--------------LeuCysLeuLysArg 203
Qy   25 GluProGluLysLeuGlySerValValIleValSerArgHisGlyValArgAlaPro 44
Db   4228 CAGCTGCGCGGGTGCAGAATGTGCCCTCACGTGACAAGGACACCTACGTCTGCATGGACCGC 4287
Db   1330 ----------CTCAGAGTTCACTCACTGGAGCTACACACAGAGTCTCAGCTTGGACCTACGGG 1292

Qy   204 GluLysGlnAspGluSerCysSerLeuThrGlnAlaAlaLeuProSerGluLysValSer 223
Qy   45 ThrLysAlaThrGlnHisLeuMetGlnAsnValThrProAspAlaTrpProThrTrpProVal 64
Db   4288 AACTCCTTGGGCTCAGTCACACCATGGGGCAGCTGCAGCCAGAAGGTGTGCACGTGGTA 4347
Db   1292 -- 1292

Qy   224 AlaAsp------------AsnValSerLeuThrGlyAlaValAlaSerLeuAlaSerMetLeuThr 240
Qy   65 LysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAla 78
Db   4348 GCTGAGCCAGGGCTGGATGTGCCTGAAGGCACAGCGCGCTGAACCTGAGCTGTCCGCCTCCCT 4407
Db   1291 AAACTTGGGGGTTTCACAGAACTGACGCGGCATTCTGGATGTTCACTGAGAGCAGCCTAG 1232

Qy   241 GluIlePheLeuLeuLeuGlnIleAlaGlnGlyMetProGluProGly------------- 255
Qy   79 ------------TyrLeuGlyHisTyrArgGlnArgGlyLeuValAlaAspGlyLeu 93
Db   4408 AGT------------------------GGCCCTGGGCACCATAGGCAACTCC 4434
Db   1231 GTGATTCACGGGTAATACTTGGAGCTACACAAATACAGGGGTTT----------CTC 1184

Qy   256 -------Trp--------GlyArgIleThrAspSerHisGlnTrpAsnThrLeuLeu 269
Qy   94 LeuAlaLysLysGlyCysProGlnSerGlyGlnValAlaIleIleIleAlaAspValAspGlu 113
Db   4435 ACCTTTGCTTGGTTCCGGAACGGTCGGCAGCTACACACAGAGTCTGTCCCACCCTTACC 4494
Db   1183 AAATCACCTAAAGGGCTCTTCAAAATAAGGATGTGTTACTCTCCAAAGAC---------- 1133

Qy   270 SerLeuHisAsnAlaGln---------------PheTyrLeuLeuGlnArgThrProGlu 284
Qy   114 ArgThrArgLysThrGlyIleGlyAlaAlaAsnValThrAspAlaIleLeuSerArgAlaGlyAlaIleThr 133
Db   4495 TTCACCCATGTGCGCCGCGCCCAGCTGGCTTGTACCCAGGCTGAGCTCCCGCC 4554
Db   1132 ------------RAGGATGCTCAGACTCGAGCGCTGGAGCTACGAGCACCCCTTCGGGCCCTGGG 1055

Qy   285 ValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAlaAlaLeuThrProHis 304
Qy   133 rValHisThrGlnAlaAspThrSerSerProAspProLeuPheAsnProLeuuLysThrGl 153
Db   4555 GGGGGCTGCCACCTCTGCTCCAGTCTTGCTCCGGGTGCTC--------------TAC 4596
Db   1054 GTCGGAGAAGAGACCCCAGCACCAAGAGCGTGCCCAAGCTTCCAGCTTCGCGCTCAGCTGT-GA 996

Qy   305 ProProGlnLysGlyAlaTyrArgGlyValThrArgLeuProThrSerValLeuPheIleAlaGly 324
Qy   153 yValCysGlnLeuLeuAspAsnAlaAsnValThrArgAspAlaIleLeuSerArgAlaGlyGlySe 173
Db   4597 CCTCCCAG--------------ACGCCACCATGACTGTTTTTGTGGAGCCC 4635
Db   1055 TGTCGCCACGGGACAGGAGCAGGACAC---AGCCCATGTGGCGCGACTGCCAAGAGCTGCTGCT 939

Qy   325 HisAspThrAsnLeuAlaAsnLeuGlyAlaAlaLeuGlu--LeuAsnTrpThrLeu---P 343
Qy   173 rIleAlaAspPheThrGlyHisArgGlnThrArgAlaPheArgGlnThrArgAlaLeuAsnValThrLeuAs 193
Db   4636 GAG--------------GGTGGCATCCAGGCATTCTGGACTGCCGAGTGGACAGTGAGCCC 4683
Db   938 TCAGCCCAGTAAACACAGA--GTTGCCTCCACGACCCCTCTTCCTGTGTCTCGCTGCTGCA 882

Qy   343 roGlyGlnProAspAsnThrProPro-GlyGlyLeuValPheGluArgTrpArgArg 362
Qy   213 rGlnAlaLeuProSerGluLeuLysValSerAlaAspAsnValSerLeuAlaSerLeuThrGlyAlaVa 233
Db   4684 CTAGCCAGCCTGAC---CCTCCACCTGGCAGTCGGCTGGTGGCCTCCAGCCAGCCTCAG 4740
Db   881 CGGAGCTTGTGGACAGGAGCTCCGCCTTGCAGCTCCTGCAGTCCAGCTCCTGGGCCACGT 822

Qy   363 LeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGlnGlnMet 382
Qy   233 lSerLeuAlaSerMetLeuThrGluIlePheLeuLeuLeuGlnIleAlaGlnGlyMet-ProG 253
Db   4741 GCTGCCCCTGCCAAGCCGCACATCCGGCGTCTCA------------ 4773
Db   821 GGACCTCAGAGCCCAAAGTGCTCAATA----------CCTGCCCCGCACCTGTCCAG 774

Qy   383 ArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThrLeuAla 402
Qy   253 luProGlyTrpGlyArgIleThrAspSerHisGlnTrpAsnThrLeuLeuSerLeuHisA 273
Db   4774 ---------GCCAGTCCAATGCCTTGCGAGTGGGACATGGAGGAGCTGAAGCCCAGTGACCAG 4827
Db   773 AACCGAAACCTGAGGCTTGACAGCTTCTTACCAGAGAGTTCCACGTGTCCCATTCCT 714

Qy   403 Gly------------CysGluGluArgAsnAlaGlnGlyMetCysSerLeuAla---------- 416
Db   4828 GGGGAGTATGTGTGCTCCGGGCTCCCAAATGCCCTGGCGCTCCTGCCTCCTGCTGCCACCTACTTC 4887

Qy   417 GlyPheThrGlnIleValAlaAsnGluAlaArgIleProAlaCysSerLeu 432
Db   4888 GGAACCCAGAGCCCCTGCA-TCGCCTGCCATCCGTTCCAGCACCTTCTCTG 4934

RESULT 6
US-11-293-697-1957/c
; Sequence 1957, Application US/11293697
; Publication No. US2006010537 6A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28

```
Qy    329  uAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrpThrLeuProGlyGlnProAsnTh  349
           :::  |||        ||| |||||||||||||  |||             |||
Db   1234  GGCCCTCCTTCTCAGGGCGAGAGATGGTGGGGCCCAGCTGCCCGGATACCACCCACAT  1293

Qy    349  rProProGlyGlyGlu  354
           ||||||
Db   1294  CCCCACCAGCGGACAG  1309


RESULT 5
US-11-270-040-5
; Sequence 5, Application US/11270040
; Publication No. US20060110762A1
; GENERAL INFORMATION:
;   APPLICANT: Kapil, Sanjay
;   APPLICANT: Kim, Jeong-Ki
;   TITLE OF INVENTION: PORCINE REPRODUCTIVE AND RESPIRATORY
;   TITLE OF INVENTION: SYNDROME VIRUS RECEPTOR COMPONENTS AND USES THEREOF
;   FILE REFERENCE: 14337/004001
;   CURRENT APPLICATION NUMBER: US/11/270,040
;   CURRENT FILING DATE:  2005-11-09
;   PRIOR APPLICATION NUMBER: 60/680,297
;   PRIOR FILING DATE:  2005-05-12
;   PRIOR APPLICATION NUMBER: 60/626,788
;   PRIOR FILING DATE:  2004-11-10
;   NUMBER OF SEQ ID NOS: 8
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 5193
;   TYPE: DNA
;   ORGANISM: Sus scrofa
US-11-270-040-5

Alignment Scores:
Pred. No.:               1.27
Score:                   109.50
Percent Similarity:      35.2%          Length:          5193
Best Local Similarity:   22.9%          Matches:         105
Query Match:                            Conservative:    56
DB:                                     Mismatches:      190
                                        Indels:          108
                                        Gaps:            21

US-10-601-319-2 (1-440) x US-11-270-040-5 (1-5193)

Qy     22  AlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGlyVal  41
           |||  :::            :::||||| ::|||||
Db   3748  GCCCAACACATCCCTGAGCTCGAGCTGCGGGGGGTGCGCGGCTAGAGGGGGTCCATCGCC  3807

Qy     42  ArgAlaProThrLysMetGlnAspValThrProAspAlaTrpProThr  61
           |||      :::
Db   3808  ACTGTCGCGGAGGGGGCCCCTCGTGTGAAGACCTGGAAGACCCTGCTGCCCCCCCA---  3864

Qy     62  TrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuLeuAlaAlaTyrLeuGly  81
           |||            :::|||
Db   3865  ---CCCACTCTTCTATGTCTGG------  3882

Qy     82  HisTyrGlnArgGlnArgLeuValAlaAspLeuLeuAlaAlaLysGlyProGln  101
           |||    :::      :::|||
Db   3883  ---TACCACACAGCCGTTGGCTGGCAGGAGGGGTCGGCTCCTCCCTCCGTTTCCA---  3936

Qy    102  SerGlyValAlaIleIleAlaAspValAspValThrArgGlyThrGlyGluAla  121
           |||      :::  |||
Db   3937  --------GCGGCTACACGGGCGCTCACGCGGCGCCTATACCTGCCCAGGATCCAGGATGCC  3987

Qy    122  PheAlaAlaGlyLeuLeuAlaProAspCysAlaIleThrVal---HisThrGlnAlaAsp---  139
           |||      :::
Db   3988  CAGGGGCCACAGCCATCTCCCAGCCCGCCACTGCACATCCTCTATGCCCCTCGGGATGCT  4047

Qy    140  ---ThrSerSerProAspProLeuPheAsnProLeuLysThrGlnAla  157
           :::   :::|||
Db   4048  GTCCTTTCCTCCTTCTCGGGACTCAAGGGCCCAGCCCTGGTCCCATGGCCGGCTACAGTG  4107

Qy    158  AspAsnAlaAsnValThrArgAspAlaAspLeuSerArgAlaGlyGlyIleAlaAspPhe  177
           |||      :::  :::|||
Db   4108  GACAGCGAGCCACCTGCCGAGATGACCCTGTCCCATGATGGCAAGGTGCTGGCCACCAGC  4167
```

```
Qy     26  ProGluLeuLysLeuGluSerValValIleValSerArgHisGlyValArgAlaAlaProThr  45
           ||||| ::: ::: :||| |:::|||
Db    239  CCGGAAGTGGTCCGCCAGCCGCATCGTAGACCTGGCCCCACCAGGTGTAAGGCCCTGCGAC  298

Qy     46  LysAlaThrGlnLeu---MetGlnAspValThrProAspAlaTrp  59
           |||||      :::
Db    299  ATCTCTCGCCAGCTCCGCGTCAGCCATGGTCGTCAGCAAGATCCTTGGCAGGTACTAC  358

Qy     60  ProThrTrpProValLysGlyLeuGlyTrpLeu---ThrPro  71
           |||                :::|||
Db    359  GAGACTGGCAGCATCCGGCCTGGAGTGATAGGGGGTCCAGCCCAAGGTGGCCACCCCC  418

Qy     72  ArgGlyGlyGluLeuLeuAlaAlaTyrLeuGlyHisTyrGlnArgGlnArg  87
           ::: :||||| |:::|||
Db    419  AAG-------GTGGTGGAGAAGATTGGGGACTACAAACGCCAGAACCCTACCATGTTT  469

Qy     88  ---LeuValAlaAspGlyLeuLeuLysLysGlyCysProGlnSerGlyLeuGlyVal  105
           ::: ||| ||||||| |||||||||
Db    470  GCCTGGGAGATCCGAGACCGGCTCCTGGCTGAGGGGCTCGT--GACAATGACACTGTG  526

Qy    106  AlaIleIleAlaAlaAspValAspGluArgThrArg---LysThrGlyGluAlaPheAlaAla  124
           ::: :::::: |||
Db    527  CCCAAGTGTCAGTCCATTAAATAGAATCATCCGGACCAAAGTGCAGCAACCATTC------  580

Qy    125  GlyLeuAlaProAspCysAlaIleThrVal---HisThrGlnAla  138
           |||                      :::|||
Db    581  AACCTCCCTATGGACAGCTGCGTGGCCACCAAGTCCCTGAGTCCCGGACACGCTGATC  640

Qy    139  AspThrSerSerProAspProLeuPheAsnProLeuLysThrGlyThrValCysGlnLeuAsp  158
           :::|||||: ::|||
Db    641  CCCAGCTCAGCTGTAACTCCCCCGGAGTCACCC------CAGTCGGAT  682

Qy    159  AsnAlaAsnValThrArgAspAlaIle-Leu---SerArgAlaGlyGlySerIleAlaAspPh  177
           ||| ::: ::: |||
Db    683  TCCCTGGCTCCACCTACTCCATCAATGGGCTTCCTGGCATCGTCAGCCTGGCAGCGAC  742

Qy    177  eThrGly---------------------HisArgGl  182
           |||||                  :::|||
Db    743  AAGAGGAAAAATGGATGACAGTGATCAGGATAGCTGCCGACTAAGCATTGACTCACAGAGC  802

Qy    182  nThrAlaPheArgGluLeuGlyLeuArgGlnLeuLeuAsnPheProGlnSerAsnLeuCysLeuGly  202
           |||                        :::::::
Db    803  AGCAGCAGGCGGACCCCGAAAGCACCTTCGACGGATGCCTTCAGCCAGCACCACCTCGAG  862

Qy    202  sArgGluLysGlnAspGluSerCysGlnThrGln---AlaLeuProSerGluLeuLeuLy  221
           |||              :::|||
Db    863  CCGCTCGAGTGCCCATTTGAGCGGCAGCACTACCCCAGCCTATGCCTCCCCAGCCAC  922

Qy    221  sValSerAlaAspAsnValSerLeuThrGlyAlaValAlaSerLeuAlaSerMetLeuThrGl  241
           |||      :::  :::|||
Db    923  ACCAAAGGGCAGCAGGGGCCCTCCACCCGCCCTTGCTCAACAGCACCCCTGACGACGGGG  982

Qy    241  uIlePheLeuLeuGlnIleAsnAlaGlnGlyMetProGluProGlyTyrTrpGlyArgIleThrAs  261
           |||                       :::::: |||
Db    983  A----------AGCCACCCTGACCCCTTCCAACACGCACCCTGGGGGCCGCAACCTCTCG  1030

Qy    261  pSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGlnAsnAr  281
           ||| :::::: :::|||
Db   1031  -ACTCACCAGACCTACCCCCGGTGGTGGCAGCACTACCCTCACTCCACCCTTGGCCATAAAGCAGGA  1089

Qy    281  gThrProGluValAlaArgSerArgAlaThrPro---------------LeuLeu  294
           |||                :::|||
Db   1090  AACCCCCGAGGGTGTCCAGTTCTAGCTCCACCCCTTTGCTCTTTATCTCCGCCCTTTT  1149

Qy    294  uAspLeuIleMetAlaAlaLeuThrProHisProHisProGlnAlaTyrGlyValTyrValTh  314
           |||                      :::|||
Db   1150  GGATCTG-------------------CAGCAAGTCGGGCTCCGG  1173

Qy    314  rLeuProThrSerValLeuPhe----------IleAlaGlyHisAspThrAsnLeu  329
           :::|||
Db   1174  GGTCCCGCCCTTCAATGCCTTTCCCCATGCTGCCCGTGCTCCGTGTACGGGCAGGGGCCA  1233
```

```
Qy   187 uLeuLeuArgValLeuAsnPheProGlnSerAsnLeu---CysLeuLysArgGluLysGl 206
             :::|||||||       ::|||       |||::: |||
Db   244 CTTGGCCGCCGCCACCAGCTTCATCGCCTCCGTGTTGCGCCAC-CGAGT 186

Qy   206 nAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeuLysValSerAlaAspAs 226
             :::           |||||||||
Db   185 CGATGCGGCTCCGAGCTCACGCAGGAGCACCGCCACCTGCAGGGCGGCCCGTCTGCTGC 126

Qy   226 nValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThrGluIlePheLeuLeuGl 246
             :::  |||||||       :::
Db   125 CGCAGCTAGGAGCGGAGCCGGCGG----------------------------------- 101

Qy   246 nGlnAlaGlnGlyMetProGluProGlyTrpProGlyTyrArgIleThrAspSerHisGlnTrp 265
             |||  :::
Db   100 -GGTGGCCAGGGCGGAGGAGGACCAGGGCGGTGGAGAGGTGCGAGCACGACATGGCTGG 44
```

RESULT 3
US-10-449-902-7723/c
; Sequence 7723, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7723
; LENGTH: 2822
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK105071
; DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-7723

Alignment Scores:
Pred. No.:   0.18   Length: 2822
Score: 114.50   Matches: 83
Percent Similarity: 37.4%   Conservative: 30
Best Local Similarity: 27.5%   Mismatches: 108
Query Match: Indels: 82
DB: Gaps: 16

US-10-601-319-2 (1-440) x US-10-449-902-7723 (1-2822)

```
Qy    12 LeuIleProLeuThrProGlnSerAlaPheAlaGlnSerGluProGluLeuLysLeuGlu 31
            |||::|||||||                                          :::
Db  2306 CTTCTCCCGCTCCACGGTCAGCTTCCCTTCCTTGGTGGTGAGCGGGA--------CAG 2256

Qy    32 SerValVal-IleValSerArgHisGlyValValArgAlaProThrLysAlaThrGlnLeuMe 51
            |||::||:|||||||              ::|||
Db  2255 CTCGTCCTCGGTGGCGTCCACGCAC-----------------ACGCCGGTTGAT 2220

Qy    51 tGlnAspValThr-ProAspAlaTrpProThrTrpProValLysLeuGly---------- 67
            :::::::|||||||                |||
Db  2219 GTCGCAGATCTCCCCT----TGGGCACATGGGACGCAGCGTCTGGATGATGATGGGGTC 2166

Qy    68 ----------------TrpLeuThrProArgGlyGly-GluLeuIleAlaTyr 79
                             |||:::               ||:::
Db  2165 GGAGCCGGACGAGGAGGACGAACTTGG-----AGTAGAGGACGTCCACCTTGTCCACCGC 2112

Qy    80 LeuGlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyGly--- 98
            |||||||       |||||||       |||||||
```

RESULT 4
US-11-145-307A-189
; Sequence 189, Application US/11145307A
; Publication No. US20060094035A1
; GENERAL INFORMATION:
; APPLICANT: Arcturus Bioscience, Inc.
; APPLICANT: Erlander, Mark G.
; APPLICANT: Ma, Xiao-Jun
; TITLE OF INVENTION: Identification of Tumors
; FILE REFERENCE: 022041-002020US
; CURRENT APPLICATION NUMBER: US/11/145,307A
; CURRENT FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: US 60/577,084
; PRIOR FILING DATE: 2004-06-04
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 189
; LENGTH: 2711
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-145-307A-189

Alignment Scores:
Pred. No.:   0.211   Length: 2711
Score: 113.50   Matches: 89
Percent Similarity: 36.8%   Conservative: 53
Best Local Similarity: 23.1%   Mismatches: 159
Query Match: 4.9%   Indels: 86
DB: Gaps: 7

US-10-601-319-2 (1-440) x US-11-145-307A-189 (1-2711)

```
Db  2111 CTCCGGACACGAACAGCGAGTAGACGAGGTCGCAGATGGACTGGGAGTCCTTCACGGTGGG 2052

Qy    99 ------------------------CysProGlnSerGlyGln---------- 104
Db    99 ------------------------CysProGlnSerGlyGln---------- 104

Qy   105 ValAlaIleIleAlaAlaAspValAspGluArgThrArgLysThrGly-----GluAlaPhe 122
            |||||:::|::::||||||:::|||
Db  1991 GTTGCCCTTCTTGCCGACGGTACGGTGGGTCTCGAGGTGCTGAGCTCCTC 1932

Qy   123 AlaAlaGlyLeuAla------ProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
            |||       :::|||                      ::|||
Db  1931 GATGCGGGTCTCGGCCTTCTTGAGGACGTTGTTGTTGAAGCTGCCGCAGAGGCCGCGCTC 1872

Qy   141 SerSerProAsp---------ProLeuPheAsnProLeuLysThrGly---------- 153
            :::                    |||
Db  1871 GCCGGTGAGGACGACGAGGGCGACCTTCTTGACGGGCGGTGCGGGTGAGGGGGAGGTC 1812

Qy   154 --ValCysGlnLeuAsp---AsnAlaAsnValThrArgAlaAlaIleLeuSerArg------ 169
            ::|||||||              ::: :::
Db  1811 GATGTCCTCCGTCTGGATCTTCCGGTTCATGTTGTAGAGCACCTCCACCAGCGCCTCCGA 1752

Qy   170 ------AlaGlySerIleAlaAspPheThrGlyHisArg-GlnThrAlaPheArgGl 187
            |||||||            |||||||
Db  1751 GAAGGGGCGGGAGGACAC----------CACCGCCTCCTGCGCCCGGCCGCCGCAC 1710

Qy   187 uLeuGluArgValLeuAsnPheProGlnSerAsnLeu---CysLeuLysArgGluLysGl 206
            :::|||||       :::|||
Db  1709 CTTGGCCGCCGCCACCAGCTTCATCGCCTCCGTGATCTTCTGCGGTGTTGCGCAC-CGAGT 1651

Qy   206 nAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeuLysValSerAlaAspAs 226
            :::           |||||||::: :::
Db  1650 CGATGCGGCTCCGAGCTCACGCAGGAGCACCGCCACCTGCAGGGCGGCCCGTCTGC 1591

Qy   226 nValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThrGluIlePheLeuLeuGl 246
            :::|||||||       :::
Db  1590 CGCAGCTAGGAGCGGAGCCGGCGGCG----------------------------- 1566

Qy   246 nGlnAlaGlnGlyMetProGluProGlyTrpProGlyTyrArgIleThrAspSerHisGlnTrp 265
            |||  :::
Db  1565 -GGTGGCCAGGGCGGAGGAGGACCAGGGCGGTGGAGAGGTGCGAGCACGACATGGCTGG 1509
```

; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 21879
; LENGTH: 1470
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK072104
; DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-21879

US-10-601-319-2 (1-440) x US-10-449-902-21879 (1-1470)

Alignment Scores:
Pred. No.:
Score: 0.0712   114.50
Percent Similarity: 37.4%
Best Local Similarity: 27.5%
Query Match: 5.0%
DB: 6

Length: 1470
Matches: 83
Conservative: 30
Mismatches: 108
Indels: 82
Gaps: 16

US-10-601-319-2 (1-440) x US-10-449-902-21879 (1-1470)

```
Qy   12 LeuIleProLeuThrProGlnSerAlaPheAlaLeuGlnSerGluProGluLeuLysLeuGlu 31
Db  841 CTTCTCCCGCTCCACCGGTCAGCTTCCCTTCCTTTCCTTGGTGGTGAGCCGGAA------CAG 791

Qy   32 SerValVal-IleValSerArgHisGlyValArgGlaAlaProThrLysAlaThrGlnLeuMe 51
Db  790 CTCGTCGTCCTCGGGTGGCGTTCCACGGCAC------------------------ACGCCGTTGAT 755

Qy   51 tGlnAspValThr-ProAspAlaTrpProThrTrpProArgLeuLeuAlaAspGlyLeuLeuAla 67
Db  754 GTCGCAGATCTCCCCCT----TGGGCGACATGGGGGAGCACGCGTCTGGATGGGGGTC 701

Qy   68 -------------TrpLeuThrProArgGlyGly-GluLeuIleAlaLeuTyr 79
Db  700 GGAGCCGGACGGAGGAGACGGAACTTGG------AGTAGAGGAGCTCCACCTTGTTCCACCGC 647

Qy   80 LeuGlyHisTyrGlnArgGlnThrArgLeuLeuAlaAspGlyLeuLeuAlaLysLysGly-- 98
Db  646 CTCCGGACCACGAACACGGCAGTGGTGTGTTCATGGGGGGCTCCAGAACGGTGGTGGGG 587

Qy   99 ------------------CysProGlnSerGlyGln--- 104
Db  586 GATGCCGTTCACCTCCAGCGTCCGGTCGGGGGATGAAGGGGGGATGAAGTAGGC 527

Qy  105 ValAlaIleIleAlaAspGluArgThrArgLysThrGly------GluAlaPhe 122
Db  526 GTTGCCCTTCTTGCCCAGACGCTGACGACGTGTACTCGAGTCGGGTGCTTGAGCTTCTC 467

Qy  123 AlaAlaGlyTyrLeuAla------ProAspCysAlaIleThrArgValHisIleThrLeuAspThr 140
Db  466 GATGCGGGGTCGTCCGGCCTTCTTGAGGACGGTTGTGTTGGAGTCGCCGGACCGGCTC 407

Qy  141 SerSerProAsp-----ProLeuPheAsnProLeuLysThrGly------------ 153
Db  406 GCCGGTGAGGACGACGAGGGCGGACCTTCTTGACGGCGGGGATGCGGGGTTGAGGGGGAGGTC 347

Qy  154 --ValCysGlnLeuAsp---AsnAlaAsnValThrArgAlaIleLeuSerArg------ 169
Db  346 GATGTCCTCGGCTGGATCTCCTGGTTCATCGTTGTTGAGAGCACTCCACCAGGCTCGGA 287

Qy  170 -----AlaGlySerIleAlaAspPheThrGlyHisArg-GlnThrAlaPheArgGl 187
Db  286 GAAGGGGGCGGGAGGACAC-------------CACCGCCTTCTCGCGCGCGCGCGCGGCAC 245
```

US-10-601-319-2 (1-440) x US-10-449-902-14217 (1-1482)

```
Qy   19 SerAlaPheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArg 38
Db  957 TCGATCTTCACCTTCTCCCGCTCCACCGGTCAGCTTCCCTTCCTTTCCTTGGTGGTGAGCCGGAAC 898

Qy   39 HisGlyValArgGlaAlaProThrLysAlaThrGlnLeuMetGlnAspValThr-ProAspAl 58
Db  897 AGCTCCGTCCTCGGGTGGCGTTCCACGGCACACGCCGTTGATGTCGCAGATCTCCCCCT----- 843

Qy   58 aTrpProThrTrpProArgValLysLeuGly-------------- 67
Db  842 -TGGGCGACATGGGGGAGCACGCGTCTGGATGATGGGGGTCGGAGCCGGACGGAGGAGAAC 784

Qy   68 -TrpLeuThrProArgGlyGly-GluLeuIleAlaLeuTyrLeuGlyHisTyrGlnArgGlnA 87
Db  783 TTGG------AGTAGAGGAGCTCCACCTTGTTCCACCGCCTCCGGACCACGAACACGGCAGTAG 730

Qy  100 roGlnSerGlyGln------------ValAlaIleIleAlaAspGluArgThr 112
Db  729 ACGAGCGTACTCGGAGTCGGGTGCTTCATGGGGGGCTCCAGAACGGTGGTGGGGGATGCCGTC 670

Qy  112 spGluArgThrArgGlyThrGly-------GluAlaPheAlaAlaGlyTyrLeuAla------P 128
Db  669 CGCTCGGTCGGGGGATGAAGGGGGGCGGGGATGAAGTAGGCCTCCGGACGACGAAGCAGAGTAG 610

Qy  145 roLeuPheAsnProLeuLysThrGly---------ValCysGlnLeuAsp---A 159
Db  609 ACGACCGGTGACTCGGAGCGGCTTCAGCGTACCTTCGGCCTTCTTG 550

Qy  159 snAlaAsnValThrArgAlaIleLeuSerArg------AlaGlyGlySerIleA 175
Db  549 AGGACGTTGTGTTGGAGTCGCCGGACCGGCTCCACCAGCGGGGGGGGGACCACC 490

Qy  175 laAspPhe------------ThrGlyHisArgGlnThrAlaPheArgGluGluG 189
Db  489 ACCTTCTTGACGGCGGGGATGCGGGGTCGTCCGGCCTTCTTGAGGACGGTTGTGTTGGAGTCGCC 430

Qy  189 luArgValLeuAsnPheProGlnSerGluAsnLeuCysLeuLysArgGlyTyrGlyAspGlus 209
Db  429 TGGTTCATGTTGTAGAGCACTCCACCAGGCTCGGAGAAGGGGGCGGGGGAGGAGACCACC 370

Qy  209 erCysSerLeuThrGlnIleAlaLeuProSerGluLeuValSerAlaAspAsnValSerL 229
Db  369 GCCTCGTCGCGCGCGCGGCACTTGGCCGCGCGCCACCAGCCTGAGCTTCTTGCTGCTCGG 310

Qy  229 euThrGlyAlaValSerLeuAlaSerMetLeuThrGluIlePheLeuLeuGlnGlnIleAlaG 249
Db  309 TGCGTCGTTGCGCAC------CGAGTCGATGCGG 283

Qy  249 lnGlyMetProGluProGlyTyrTrpGlyArgIleThrArgAspSerHisGlnTrp 265
Db  282 CTCCGCAGCTCACGAGCGAGCACCGCCACCTGCGCTGCTCGAGCAGCTA 223

Qy  198 AGGGCCGGAGGAGGACCAGGGCGTTGGCGACGACATGGCTGG 149
Db  198 GGATCGCGAGGCGGCGGCC 199
```

RESULT 2
US-10-449-902-21879/c
; Sequence 21879, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.

Query Match:
DB:

OM protein - nucleic search, using frame_plus_p2n model

```
                GenCore version 5.1.9
        Copyright (c) 1993 - 2006 Bioacceleration Ltd.
```

Run on: June 14, 2006, 14:22:06 ; Search time 33.8073 Seconds
                                     (without alignments)
                                     4595.351 Million cell updates/sec

Title: US-10-601-319-2
Perfect score: 2302
Sequence: 1 MKAILIPFLSLIIPLTPQSA...........IVNBARIPACSLRSHHHHHH 440

Scoring table: BLOSUM62
```
           Xgapop 10.0 ,  Xgapext  0.5
           Ygapop 10.0 ,  Ygapext  0.5
           Fgapop  6.0 ,  Fgapext  7.0
           Delop   6.0 ,  Delext   7.0
```

Searched: 290227 seqs, 117694381 residues

Total number of hits satisfying chosen parameters: 580454

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
```
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10601319//runat_12062006_180105_27789/app_query.fasta_1
-DB=Published_Applications_NA_New -QFMT=fastap -SUFFIX=rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=200 -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abss03h
-USER=US10601319 @CGN_1_1_28 @runat_12062006_180105_27789
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :   Published Applications NA New:*
```
1:   /EMC_Celerra_SIDS37/ptodata/2/pubpna/US09_NEW_PUB.seq:*
2:   /EMC_Celerra_SIDS3/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3:   /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4:   /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_NEW_PUB.seq:*
5:   /EMC_Celerra_SIDS3/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
6:   /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10_NEW_PUB.seq:*
7:   /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq:*
8:   /EMC_Celerra_SIDS3/ptodata/2/pubpna/US60_NEW_PUB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	118	5.1	1482	6	US-10-449-902-14217	Sequence 14217, A
c	2	114.5	5.0	1470	6	US-10-449-902-21879	Sequence 21879, A
c	3	114.5	5.0	2822	6	US-10-449-902-7723	Sequence 7723, Ap
c	4	113.5	4.9	2711	6	US-11-145-307A-189	Sequence 189, Appl
c	5	109.5	4.8	5193	7	US-11-270-040-5	Sequence 5, Appli
c	6	101.5	4.4	1945	7	US-11-293-697-1957	Sequence 1957, Ap
c	7	101	4.4	2259	7	US-11-217-529-274	Sequence 274, Appl
c	8	99	4.3	3287	6	US-10-449-902-12801	Sequence 12801, A
c	9	98.5	4.3	14756	6	US-10-505-928-570	Sequence 570, App
	10	97	4.2	1502	7	US-11-297-134-55	Sequence 55, Appl
	11	97	4.2	1619	7	US-10-953-349-34901	Sequence 34901, A
	12	96	4.2	2010	6	US-10-449-902-20112	Sequence 20112, A
	13	95.5	4.1	2293	6	US-10-449-902-26929	Sequence 26929, A
	14	95	4.1	1488	6	US-10-449-902-24588	Sequence 24588, A
	15	95	4.1	3731	6	US-10-449-902-27838	Sequence 27838, A
	16	94.5	4.1	2329	7	US-11-293-697-1489	Sequence 1489, Ap
	17	94.5	4.1	3066	6	US-10-449-902-12659	Sequence 12659, A
	18	94.5	4.1	3773	7	US-11-293-697-841	Sequence 841, App
	19	94.5	4.1	4239	6	US-10-511-834-543	Sequence 543, App
c	20	94.5	4.1	12225	7	US-10-501-834-14	Sequence 14, Appl
	21	94	4.1	1221	1	US-11-284-978-71	Sequence 71, Appl
	22	93	4.0	1948	1	US-09-949-925-1	Sequence 1, Appli
c	23	92.5	4.0	6152	6	US-10-528-659-6	Sequence 6, Appli
	24	91.5	4.0	1720	6	US-10-449-902-20120	Sequence 20120, A
c	25	91.5	4.0	3317	7	US-11-293-697-1085	Sequence 1085, Ap
	26	91.5	4.0	3646	7	US-11-511-937-358	Sequence 358, App
	27	91	4.0	2127	6	US-10-953-349-8294	Sequence 8294, Ap
	28	91	4.0	2168	6	US-10-449-902-16137	Sequence 16137, A
c	29	91	4.0	2295	7	US-11-293-697-946	Sequence 946, App
	30	91	4.0	2970	6	US-10-449-902-1467	Sequence 1467, App
c	31	90.5	3.9	1339	7	US-11-293-697-1494	Sequence 1494, Ap
	32	90.5	3.9	1450	6	US-10-449-902-22017	Sequence 22017, A
c	33	90.5	3.9	1526	6	US-10-449-902-13489	Sequence 13489, A
	34	90	3.9	1719	6	US-10-449-902-26664	Sequence 26664, A
	35	90	3.9	1987	7	US-11-293-697-1200	Sequence 1200, Ap
	36	90	3.9	2408	6	US-10-449-902-315	Sequence 315, App
	37	90	3.9	3069	6	US-10-449-902-12653	Sequence 12653, A
	38	90	3.9	13128	6	US-11-118-524-1	Sequence 1, Appli
	39	90	3.9	138941	7	US-11-489-730-10	GENERAL INFORMATI
	40	89.5	3.9	1143	7	US-11-293-697-2750	Sequence 2750, Ap
	41	89.5	3.9	1927	7	US-10-449-902-11403	Sequence 11403, A
	42	89.5	3.9	2703	7	US-11-293-697-518	Sequence 518, App
	43	89.5	3.9	2751	7	US-11-293-697-815	Sequence 815, App
	44	89	3.9	2223	6	US-10-449-902-22201	Sequence 22201, A
	45	89	3.9	2413	6	US-10-449-902-11507	Sequence 11507, A

ALIGNMENTS

RESULT 1
US-10-449-902-14217/c
Sequence 14217, Application US/10449902
Publication No. US20060123505A1

GENERAL INFORMATION:
```
APPLICANT: National Institute of Agrobiological Sciences.
APPLICANT: Bio-oriented Technology Research Advancement Institution.
APPLICANT: The Institute of Physical and Chemical Research.
APPLICANT: Foundation for Advancement of International Science.
TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
FILE REFERENCE: MOA-A0205Y1-US
CURRENT APPLICATION NUMBER: US/10/449,902
CURRENT FILING DATE: 2003-05-29
PRIOR APPLICATION NUMBER: JP 2002-203269
PRIOR FILING DATE: 2002-05-30
PRIOR APPLICATION NUMBER: JP 2002-383870
PRIOR FILING DATE: 2002-12-11
NUMBER OF SEQ ID NOS: 56791
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 14217
  LENGTH: 1482
  TYPE: DNA
  ORGANISM: Oryza sativa
PUBLICATION INFORMATION:
  DATABASE ACCESSION NUMBER: AK064835
  DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-14217
```

Alignment Scores:
Pred. No.:

Pred: 0.0036	Length: 1482
Score: 118.00	Matches: 74
Percent Similarity: 34.9%	Conservative: 30
Best Local Similarity: 24.8%	Mismatches: 116

GenCore version 5.1.9

Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 14, 2006, 14:19:34 ; Search time 1438.58 Seconds
                                   (without alignments)
                              5637.400 Million cell updates/sec

Title: US-10-601-319-2
Perfect score: 2302
Sequence: 1 MKAILIPFLSLLIPLTPQSA.........IVNEARIPACSLRSHHHHHH 440

Scoring table: BLOSUM62
               Xgapop   10.0 , Xgapext   0.5
               Ygapop   10.0 , Ygapext   0.5
               Fgapop   6.0 , Fgapext   7.0
               Delop   6.0 , Delext   7.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/abss/ABSSWEB_spool/US10601319/runat_12062006_180103_27735/app_query.fasta_1
-DB=Published Applications NA_Main -QFMT=fastap -SUFFIX=rnpbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss06p
-USER=US10601319 @CGN_1_1_3524 @runat_12062006_180103_27735 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=30 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :   Published Applications NA_Main:*
              1: /EMC_Celerra_SIDS37/ptodata/2/pubpna/US07_PUBCOMB.seq:*
              2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:*
              3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
              4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
              5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
              6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
              7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
              8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
              9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
             10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
             11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
             12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
             13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
             14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq:*
             15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq:*
             16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result   Description
No.   Score Query Match Length DB ID
------ ----- ----- ----- ------ -- --   -----------
  1   2302   100.0 1323   3 US-09-777-566A-1   Sequence 1, Appli

No.	Score	Match	Length		ID		Description
2	2302	100.0	1323	3	US-09-866-379-1		Sequence 1, Appli
3	2302	100.0	1323	6	US-10-034-985-1		Sequence 1, Appli
4	2302	100.0	1323	7	US-10-430-356-1		Sequence 1, Appli
5	2302	100.0	1323	8	US-10-601-319-1		Sequence 1, Appli
6	2302	100.0	1323	10	US-10-933-115-1		Sequence 1, Appli
7	2235	97.1	1299	8	US-10-282-122A-7167		Sequence 7167, Ap
8	2235	97.1	1901	3	US-09-866-379-7		Sequence 7, Appli
9	2235	97.1	1901	6	US-10-156-660-3		Sequence 3, Appli
10	2235	97.1	1901	8	US-10-601-319-7		Sequence 7, Appli
11	2235	97.1	1901	10	US-10-933-115-7		Sequence 7, Appli
12	2235	97.1	1901	14	US-11-056-354-3		Sequence 3, Appli
13	2228	96.8	1901	8	US-10-601-319-6		Sequence 6, Appli
14	2226	96.7	1901	3	US-09-866-379-9		Sequence 9, Appli
15	2221	96.5	1489	3	US-10-266-041-9		Sequence 9, Appli
16	2221	96.5	1489	7	US-10-284-962-1		Sequence 1, Appli
17	2221	96.5	1901	3	US-09-866-379-6		Sequence 6, Appli
18	2221	96.5	1901	8	US-10-601-319-5		Sequence 5, Appli
19	2221	96.5	1901	10	US-10-933-115-5		Sequence 5, Appli
20	2221	96.5	1901	10	US-10-933-115-6		Sequence 6, Appli
21	2220	96.4	1901	3	US-09-866-379-5		Sequence 5, Appli
22	2217	96.3	1486	13	US-11-018-709-2		Sequence 2, Appli
23	2197	95.4	1486	7	US-10-284-962-4		Sequence 4, Appli
24	2197	95.4	1486	13	US-11-018-709-4		Sequence 4, Appli
25	2188	95.0	1308	6	US-10-156-660-1		Sequence 1, Appli
26	2188	95.0	1308	8	US-10-601-319-9		Sequence 9, Appli
27	2188	95.0	1308	10	US-10-933-115-9		Sequence 9, Appli
28	2171	94.3	1296	14	US-11-056-354-1		Sequence 14, Appl
29	2081.5	90.4	11357	13	US-11-074-522-14		Sequence 14, Appl
30	2070	89.9	1281	7	US-10-334-672-4		Sequence 4, Appli
31	2070	89.9	1281	7	US-10-334-671-4		Sequence 4, Appli
32	1737	52.0	1737	10	US-10-450-763-25689		Sequence 25689, A
33	1196.5	41.7	1326	8	US-10-021-723A-11		Sequence 11, Appl
34	960.5	41.0	1325	8	US-10-282-122A-41608		Sequence 41608, A
35	943.5	40.8	1325	6	US-10-021-723A-3		Sequence 3, Appli
36	938.5	39.9	1431	6	US-10-021-723A-1		Sequence 1, Appli
37	918.5	33.0	4050	10	US-10-450-763-386		Sequence 386, App
38	760	30.8	466	10	US-10-282-122-3181		Sequence 3181, Ap
39	759.5	28.5	1264	10	US-10-450-763-24247		Sequence 24247, A
40	708	23.4	1264	10	US-10-450-763-30120		Sequence 30120, A
41	657	20.7	11710	7	US-10-240-689-40		Sequence 40, Appl
42	657	20.5	539	6	US-10-021-723A-7		Sequence 7, Appli
43	539	17.9	1230	6	US-10-021-723A-5		Sequence 5, Appli
44	476.5		1266	16	US-11-098-686-9338		Sequence 9338, Ap
45	472.5		1326				
	412.5						

ALIGNMENTS

RESULT 1
US-09-777-566A-1

; Sequence 1, Application US/09777566A
; Patent No. US2001005788A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-6
; CURRENT APPLICATION NUMBER: US/09/777,566A
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 1323
;   TYPE: DNA

```
;   ORGANISM: Escherichia coli
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1320)
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1323)
;   OTHER INFORMATION: n is any nucleotide
US-09-777-566A-1

Alignment Scores:
Pred. No.:                   1.32e-261        Length:           1323
Score:                       2302.00          Matches:          440
Percent Similarity:          100.0%           Conservative:     0
Best Local Similarity:       100.0%           Mismatches:       0
Query Match:                 100.0%           Indels:           0
DB:                          3                Gaps:             0

US-10-601-319-2 (1-440) x US-09-777-566A-1 (1-1323)

Qy     1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla    20
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA    60

Qy    21 PheAlaGlnSerGluProGluLeuLysGluGluSerValValIleValSerArgHisGly    40
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TTCGCTCCAGTGAGCCGGAGCTGAAAGCTGAAAGTGTGGTGATTGTCAGTCGTCATGGT   120

Qy    41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro    60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA   180

Qy    61 ThrTrpProValLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu    80
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ACCTGGCCGGTAAACTGGGTTGGCTGACACCGCGNGGTGTGGAGCTAATCGCCTATCTC   240

Qy    81 GlyHisTyrArgGlnArgGlyLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro   100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG   300

Qy   101 GlnSerGlyGlnValAlaIleIleAlaAspValAspLeuArgThrArgLysThrGlyGlu   120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA   360

Qy   121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr   140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCCTTCGCCGCGGGCTGGCAACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG   420

Qy   141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla   160
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCCAGTCCGGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG   480

Qy   161 AsnValThrAspAlaAlaLeuSerArgAlaAlaGlyGlySerIleAlaAspPhePheGlyHis   180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AACGTGACTGACGCGGCATCCTCAGCGAGGGCGCAATTGCTGACTTTTACCGGGCAT     540

Qy   181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys   200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC   600

Qy   201 LeuLysArgGluLysGlnAspSerCysSerLeuThrGlnLeuProSerGluLeu        220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTTAAACGTGAGAAACAGGACCAAAGCTGTTCATTAACCGCAGCATTACCATCGGAACTC   660

Qy   221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerAlaSerMetLeuThr     240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCGCCTCGCATCAATGCTGACG   720

Qy   241 GluIlePheLeuLeuGlnAsnAlaGlnGlyMetProGluGlyGlyTrpGlyArgIleThr   260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GAGATATTTTCTCCTGCAACAAGCACACAGGGAATGCCCGGAGTGGGGAAGGATCACC     780

Qy   261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln   280
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   781 GATTCACCACCAGTGGAACACCTTGCTAAGTTGCATAACGCGCAATTTATTGCTACAA   840

Qy   281 ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuAspLeuIleMetAlaAla   300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CGCACGCCAGAGGTTGCCCGCAGCCGCGGCCACCCCGGTTATTGGATTTGATCATGGCACG   900

Qy   301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu   320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTGACGGCCCCATCCACCACGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTACTG   960

Qy   321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp     340
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TTTATTGCGGCACGACGATACTAATCTGCCAAATCTCCGCGGGCACTGGAGCTCAACTGG   1020

Qy   341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp   360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 ACGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGTGGTGAACTGGTTTGAACGCTGG   1080

Qy   361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnInValSerLeuValPheGlnInThrLeuGln   380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 CGTCGGCTAAGGCGATAACAGCCAGTGGATTCAGGTTTCGTCTGCTCTTCCAGACTTTACAG   1140

Qy   381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyValValLysLeuThr   400
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCCGGAGAGGTGAAACTGACC   1200

Qy   401 LeuAlaGlyCysGluGluValArgAsnAlaGlnAsnGlyMetCysSerLeuAlaGlyPheThrGln   420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 CTGGCAGGATGTGAAGAGGCGAAATGCCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA   1260

Qy   421 IleValAsnGluAlaArgIleProAlaCysSerSerLeuArgSerHisHisHisHisHis   440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTGAGATCTCATCATCACCACCATCAC   1320
```

```
RESULT 2
US-09-866-379-1
;   Sequence 1, Application US/09866379
;   Patent No. US20020136754A1
;   GENERAL INFORMATION:
;   APPLICANT: DIVERSA CORPORATION
;   APPLICANT: SHORT, Jay
;   APPLICANT: KRETZ, Keith
;   APPLICANT: GRAY, Kevin
;   APPLICANT: BARTON, Nelson
;   APPLICANT: GARRETT, James
;   APPLICANT: O'DONOGHUE, Eileen
;   TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
;   FILE REFERENCE: DIVER1370-7
;   CURRENT APPLICATION NUMBER: US/09/866,379
;   CURRENT FILING DATE:  2001-05-24
;   PRIOR APPLICATION NUMBER: US 09/580,515
;   PRIOR FILING DATE: 2000-05-25
;   PRIOR APPLICATION NUMBER: US 09/318,528
;   PRIOR FILING DATE: 1999-05-25
;   PRIOR APPLICATION NUMBER: US 09/291,931
;   PRIOR FILING DATE: 1999-04-13
;   PRIOR APPLICATION NUMBER: US 09/259,214
;   PRIOR FILING DATE: 1999-03-01
;   PRIOR APPLICATION NUMBER: US 08/910,798
;   PRIOR FILING DATE: 1997-08-13
;   NUMBER OF SEQ ID NOS: 10
;   SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
;   LENGTH: 1323
;   TYPE: DNA
;   ORGANISM: Escherichia coli
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1323)
;   OTHER INFORMATION: n is any nucleotide
;   NAME/KEY: CDS
;   LOCATION: (1)..(1323)
;   OTHER INFORMATION:
```

US-09-866-379-1

Alignment Scores:
Pred. No.:
Score: 1.32e-261
Percent Similarity: 2302.00
Best Local Similarity: 100.0%
Query Match: 100.0%
DB:

Length: 1323
Matches: 440
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-10-601-319-2 (1-440) x US-09-866-379-1 (1-1323)

```
Qy    1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60

Qy   21 PheAlaGlnSerGluProGluLeuLysGluSerValValIleValSerArgHisGly 40
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCGGAGCTGAAGTGGGTCAGGCACGGT 120

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy   61 ThrTrpProValLysGlyValTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCCGCGGGTGGTGAGCTAATCGCCTATCTC 240

Qy   81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyLysCysPro 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG 300

Qy  101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGTCTGGTCAGTCGGCGATTATTGCTGACGTGGACGAGCGTACCCGTAAAACAGGCGAA 360

Qy  121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420

Qy  141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TCCAGTCCCGATCCGCTGTTTATTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG 480

Qy  161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 AACGTGACTGACGCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 540

Qy  181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600

Qy  201 LeuLysArgGluLysLysGlnAspGluSerLeuThrGlnLeuAlaGlnProSerGluLeu 220
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CTTAAACGTGAGAAAAAGCAGGACGAAAGTCTTCATTAAACGCAGGCATTACCATCGGAACTC 660

Qy  221 LysValSerAlaAspAsnAlaValSerLeuValThrGlyAlaAlaValSerLeuAlaSerMetAlaThr 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 AAGGTGAGCGCCGACAATGCTGTCATTAACCGTGCCGTAAGCCTCGGCATCAATGCTGACG 720

Qy  241 GluIlePheLeuLeuGlnGlnAlaGlnGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GAGATATTTCTCCTGCAACAAGCACAGCAAGGGATGCCGGAGCCGGGATGGGGAAGGATCACC 780

Qy  261 AspSerHisGlnThrTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA 840

Qy  281 ArgThrProHisProGlnLysGlnAlaGlnAlaTyrGlyValThrLeuProThrSerValLeu 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 CGCACGCCAGAGGTTGCCCAGCCCAGACCCGCCACCCCGTTATTGGATTGGATCATGGCAGCCG 900

Qy  301 ArgThrProHisProGlnLysGlnAlaGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
```

RESULT 3
US-10-034-985-1
; Sequence 1, Application US/10034985
; Publication No. US20030049815A1
GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/034,985
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1320)
NAME/KEY: misc_feature
LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G
US-10-034-985-1

Alignment Scores:
Pred. No.:
Score: 1.32e-261
Percent Similarity: 2302.00
Best Local Similarity: 100.0%
Query Match: 100.0%
DB:

Length: 1323
Matches: 440
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 6

US-10-601-319-2 (1-440) x US-10-034-985-1 (1-1323)

```
Qy    1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60
```

```
Db  901 TTGACGCCCATCCACCGCAAAAACAGGGTATGGTGTGACATTACCACTTCCAGTACTG 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 340
Db  961 TTTATTGCCGGACACGATCTAATCTGGCAAATCTCGGCGAGCTCCAACTCG 1020
Qy  341 ThrLeuProGlyGlnProAspAsnThrProGlyGlyGluLeuValPheGluValArgTrp 360
Db 1021 ACGCTTCCCGGTCAGCCGGATAACACGCCGGGAGGTGGTGAACTGGTGTTTGAACGCTGG 1080
Qy  361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140
Qy  381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProGlyValLysLeuThr 400
Db 1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCCGGAGAGGTGAAACTGACC 1200
Qy  401 LeuAlaGlyCysGluGluLysArgAsnAlaGlnAlaGlnAlaGlyPheThrGln 420
Db 1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA 1260
Qy  421 IleValAlaSnGluAlaArgIleProAlaCysSerSerLeuArgSerHisHisHisHisHis 440
Db 1261 ATCTGTGAATGAAGCAGCGCATACCCGTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320
```

```
Db  1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGGCCGCCCGGAGAGGTGAAACTGACC 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln   420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 CTGGCAGGATGTGAAGAGCGAAATGCCAGGGCATGTGTCCTGTGGCCAGGTTTTACGCAA 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHisHis   440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 ATCGTGAATGAAGCACGCATACCGGGCTGCAGTTTGAGATCTCATCACCACCATCAC    1320
```

RESULT 4
US-10-430-356-1
; Sequence 1, Application US/10430356
; Publication No. US2003032041A1
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/430,356
; CURRENT FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G
US-10-430-356-1

```
Alignment Scores:
Pred. No.:                                                          1323
Score:                     1.32e-261        Length:                  440
Percent Similarity:        100.0%           Matches:                   0
Best Local Similarity:     100.0%           Conservative:              0
Query Match:               7                Mismatches:                0
DB:                                         Indels:                    0
                                            Gaps:                      0
```

US-10-601-319-2 (1-440) x US-10-430-356-1 (1-1323)

```
Qy    1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla  20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCGATCTTCGTTAACCCCGCAATCTGCA  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleSerArgHisGly     40
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGGGTGATTGTCAGTCGTCATCATGGT      120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   61 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro  80
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC
```

(Left column — US-10-601-319-2 query × db)

```
Qy   21 PheAlaGlnSerGluProGluLysLeuLeuSerValValIleValSerArgHisGly    40
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGGTGATTGTCAGTCGTCATCATGGT      120

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy   61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu  80
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC 240

Qy   81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GGACATTACCAACGCCAGCGCCTGGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG 300

Qy  101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGTCTGGTCAGGTGGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA 360

Qy  121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GCCTTCGCCGCGGGTCTGGCACCGGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420

Qy  141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TCCAGTCCGGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG 480

Qy  161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 AACGTGACTGACGCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 540

Qy  181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600

Qy  201 LeuLysArgGluLysGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu        220
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CTTAAACGTGAGAAACAGGAGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC 660

Qy  221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTGGCGTCGATGCTGACG 720

Qy  241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGTTGGGGAAGGATCACC 780

Qy  261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA 840

Qy  281 ArgThrProGluValAlaArgSerArgAlaThrProLeuAspLeuLeuMetGlnAlaAla 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTGGATTTGATCATGGCAGCG 900

Qy  301 LeuThrProHisProProGlnLysAlaGlnAlaTyrArgValValThrLeuProThrVal 320
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TTGACGCCCCATCCACCGCAAAAAGCACAGGCGTATCGTGTGACATTACCCACTGTACTG 960

Qy  321 PheIleAlaGlyHisAspAspAsnValAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 TTTATTGCCGGACACGATGACAATGTCGCTAATCTCGGCGGCGCACTGGAGCTCAACTGG 1020

Qy  341 ThrLeuProGlyGlnAsnProThrProProGlyGlyGluLeuValPheGluLysTrp    360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 ACGCTTCCCGGTCAGAACCCGACGCCGCCGGGTGGTGAACTGGTGTTTGAAAAGTGGCTG 1080

Qy  361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140

Qy  381 GlnMetArgAspLysSerThrProLeuSerLeuAsnThrProProGlyGlyValLeuThr 400
```

```
; Sequence 1, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
;  APPLICANT: Short, Jay M.
;  APPLICANT: Kretz, Keith A.
;  APPLICANT: Gray, Kevin A.
;  APPLICANT: Barton, Nelson Robert
;  APPLICANT: Garrett, James B.
;  APPLICANT: O'Donoghue, Eileen
;  APPLICANT: Mathur, Eric J.
;  TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
;  TITLE OF INVENTION:  AND USING THEM
;  FILE REFERENCE: 09010-029011
;  CURRENT APPLICATION NUMBER: US/10/601,319
;  CURRENT FILING DATE:  2003-06-20
;  PRIOR APPLICATION NUMBER: US 09/866,379
;  PRIOR FILING DATE: 2001-05-24
;  PRIOR APPLICATION NUMBER: US 09/580,515
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: US 09/318,528
;  PRIOR FILING DATE: 1999-05-25
;  PRIOR APPLICATION NUMBER: US 09/291,931
;  PRIOR FILING DATE: 1999-04-13
;  PRIOR APPLICATION NUMBER: US 09/259,214
;  PRIOR FILING DATE: 1999-03-01
;  PRIOR APPLICATION NUMBER: US 08/910,798
;  PRIOR FILING DATE: 1997-08-13
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn for Windows Version 4.0
; SEQ ID NO 1
;  LENGTH: 1323
;  TYPE: DNA
;  ORGANISM: Escherichia coli
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (1)...(1320)
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: 216
;  OTHER INFORMATION: n = A,T,C or G
US-10-601-319-1

Alignment Scores:
Pred. No.:                                  Length:            1323
Score:             1.32e-261                Matches:           440
Percent Similarity:  2302.00                Conservative:      0
Best Local Similarity: 100.0%              Mismatches:        0
Query Match:       100.0%                   Indels:            0
DB:                8                        Gaps:              0

US-10-601-319-2 (1-440)  x  US-10-601-319-1 (1-1323)

Qy     1 MetLysAlaIleLeuIleProProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
Db     1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCGATTCCGTTAACCCCGCAATCTGCA 60

Qy    21 PheAlaGlnSerGluProGluLysLysLeuLeuGluSerValValIleValSerArgHisGly 40
Db    61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGGTGGTGATTGTCGTCGTCATGGT 120

Qy    41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAlaTrpPro 60
Db   121 GTGCGTGCTCCAACCAAGGCCACCAACTGATGCAGGATGTCACCCCAGCATGGCCA 180

Qy    61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db   181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCCGNGGTGGTGAGCTAATCGCCTATCTC 240

Qy    81 GlyHisTyrGlnArgGlnArgLeuValAlaGlyLeuLeuAlaLysGlyGlyCysPro 100
Db   241 GGACATTACCAACGCCACCGTCGTGAGCGACGACGATTGCTGGCGAAAAGGGCTGCCCG 300

Qy   101 GlnSerGlnValAlaIleAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db   241 GGACATTACCAACGCCACCGTCGTGGTAGCCGACGGATTGCTGGCGAAAAGGGCTGCCCG 300

Qy   101 GlnSerGlnValAlaIleAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db   241 GGACATTACCAACGCCACCGTCGTGGTAGCCGACGGATTGCTGGCGAAAAGGGCTGCCCG 300

Qy   121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db   301 GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACGTACATACCCAGGCAGATACG 420

Qy   141 SerSerProAspAspProLeuAsnProAsnPheThrGlyValCysGlnLeuAspAsnAla 160
Db   421 TCCAGTCCCGATCCGTTATTTAATCCTCAAAAACCTGGCGTTTGCCAACTGGATAACGCG 480

Qy   161 AsnValThrAspAlaIleLeuSerArgAlaAlaGlySerIleAlaAspPheThrGlyHis 180
Db   481 AACGTGACTGACGCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 540

Qy   181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db   541 CGGCAAACGGCGTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600

Qy   201 LeuLysArgGluLysGlnAspLeuSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db   601 CTTAAACGTGAGAAACAGGACGAAAAGTGTTCATTACGCAGGCATTACCATCGGAACTC 660

Qy   221 LysValSerAlaAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db   661 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCCATCAATGCTGACG 720

Qy   241 GluIleProGlyLeuIleGlyGlnAlaGlnProLysMetProGlyTrpProGlyIleArgIleThr 260
Db   721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCCGGAGCCGGGGTGGGAAGGATCACC 780

Qy   261 AspSerHisGlnArgTrpAsnThrLeuAsnSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db   781 GATTCACACCAGTGGAACACCTTGCTAAGTTGCATAACGCGCAATTTATTTGCTACAA 840

Qy   281 ArgThrProGluValAlaArgSerArgAlaThrProLeuAspLeuIleLeuMetAlaAla 300
Db   841 CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTGGATTTGATCATGATGGCAGCG 900

Qy   301 LeuThrProHisProGlnLysGlnIleAlaTyrGlyValThrLeuProThrSerValLeu 320
Db   901 TTGACGCCCCATCCACCCGCAAAAACAGGCGCTATGGTGTGACATTACCCACTTCAGTACTG 960

Qy   321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db   961 TTTATTGCCGGACACGATACTAATCTCGGCAAATCTCGGAGCTCAACTGG 1020

Qy   341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db  1021 ACGCTTCCCGGTCAGCCGGATAACACCCGCCCCGGCGGTGAACTGGTGTTTGAACGCTGG 1080

Qy   361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db  1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140

Qy   381 GlnMetArgAspLysProSerLeuSerAsnThrProProGlyGluValLysLeuThr 400
Db  1141 CAGATGCGTGATAAAAACCGCCGTGTCATTAAATACGCCCGGGAGAGGTGAAACTGACC 1200

Qy   401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuSerLeuAlaGlyThrGln 420
Db  1201 CTGGCAGGATGTGAAGAGCGAAATGCCCAGGCCATGTGCGTTGGCAGGTTTTACCCAA 1260

Qy   421 IleValAsnGluValAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
Db  1261 ATCGTGAATGAAGCAGCACACGCATACCCGGCGTCAGTTGAGATCTTCATCACCACCATCAC 1320

RESULT 5
US-10-601-319-1
```

```
;  APPLICANT:    Kretz, Keith A.
;  APPLICANT:    Gray, Kevin A.
;  APPLICANT:    Barton, Nelson Robert
;  APPLICANT:    Garrett, James B.
;  APPLICANT:    O'Donoghue, Eileen
;  APPLICANT:    Mathur, Eric J.
;  TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
;  TITLE OF INVENTION:   THEREOF
;  FILE REFERENCE: 564462001822
;  CURRENT APPLICATION NUMBER: US/10/933,115
;  CURRENT FILING DATE:   2004-09-01
;  PRIOR APPLICATION NUMBER: US/09/866,379
;  PRIOR FILING DATE:   2001-05-24
;  PRIOR APPLICATION NUMBER: US 09/580,515
;  PRIOR FILING DATE:   2000-05-25
;  PRIOR APPLICATION NUMBER: US 09/318,528
;  PRIOR FILING DATE:   1999-05-25
;  PRIOR APPLICATION NUMBER: US 09/291,931
;  PRIOR FILING DATE:   1999-04-13
;  PRIOR APPLICATION NUMBER: US 09/259,214
;  PRIOR FILING DATE:   1999-03-01
;  PRIOR APPLICATION NUMBER: US 08/910,798
;  PRIOR FILING DATE:   1997-08-13
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: FastSEQ for Windows Version 4.0
;
SEQ ID NO 1
;  LENGTH: 1323
;  TYPE: DNA
;  ORGANISM: Escherichia coli
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (1)...(1320)
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: 216
;  OTHER INFORMATION: n = A,T,C or G
US-10-933-115-1

Alignment Scores:
Pred. No.:                              Length:        1323
Score:                    1.32e-261     Matches:        440
Percent Similarity:       2302.00       Conservative:     0
Best Local Similarity:    100.0%        Mismatches:       0
Query Match:              100.0%        Indels:           0
DB:                       100.0%        Gaps:            10

US-10-601-319-2 (1-440) x US-10-933-115-1 (1-1323)

US-10-933-115-1 (1-440) x US-10-933-115-1 (1-1323)

Qy      1  MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla    20
Db      1  ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTTATACCGCTTAACCCCGCAATCTGCA    60

Qy     21  PheAlaGlnSerGluProGluLeuLysSerValValIleLeuSerValIleValSerArgHisGly    40
Db     61  TTCGCTCAGAGTGAGCCGGAGCTGAAGTGGGTTGTGTCAGTCGTCATGGT   120

Qy     41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnLeuAspValThrProAlaTrpPro    60
Db    121  GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGCATGGCCA   180

Qy     61  ThrTrpProValLysProArgGlyTrpLeuThrProArgGlyGluLeuAlaTyrLeu    80
Db    181  ACCTGGCCGGTAAAACCGCGCGGGTGGCTGACACCGCGNGGTGGTGAGCTGGCTTATCTC    240

Qy     81  GlyHisTyrGlnArgGlnAlaIleLeuAlaAspValAspGluArgThrArgLysThrGlyGlu   100
Db    241  GGACATTACCAACGCCAGGCTATATTGGCCGACGAGCGTACCCGTAAAACAGGCGAA   300

Qy    101  GlnSerGlyGlnValAlaIleLeuAlaAspCysSerAlaIleThrValHisThrGlnValHisGluLeuThrGlnAlaAspThr   120
Db    301  CAGTCTCGGTCAGGTGCCGATTATTGCTGATTGTCGAGAGCGTACCCGTAAAACAGGCGAA   360

Qy    121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr   140
```

```
Db    301  CAGTCTGGTCAGGTGCCGATTATTGCTGATTGTCGAGAGCGTACCCGTAAAACAGGCGAA    360
Qy    121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr    140
Db    361  GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG    420
Qy    141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspValAsnProAlaAla    160
Db    421  TCCAGTCCCGATCCGTATTTTAATCCTCTAAAAACTGGCGTTGCCAACTGGGATAACGCG    480
Qy    161  AsnValThrAspAlaAlaIleLeuSerArgAlaGlyAlaGlyGlySerIleLeuAlaAspPheThrGlyHis    180
Db    481  AACGTGACTGACGCCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT    540
Qy    181  ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys    200
Db    541  CGGCAAACGGCGGTTCGGAACTGGAACTGGGTGCTTAATTTTCCGCAATCAAACTGTGTGC    600
Qy    201  LeuLysArgGluLysGlnGluCysSerSerLeuThrGlnAlaLeuProSerGluLeu    220
Db    601  CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC    660
Qy    221  LysValSerAlaAlaAspAsnValSerLeuThrGlyGlyAlaValSerLeuLeuSerMetLeuThr    240
Db    661  AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG    720
Qy    241  GluIlePheLeuLeuGlnGlnIleAlaGlyLeuMetProGlyProTrpGlyValArgIleThr    260
Db    721  GAGATATTTCTTCTGCAACAACAGCACAGGATGGCCGAGCCCGGAGGTGGGGAAGGATCACC    780
Qy    261  AspSerHisGlnThrArgAsnThrLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln    280
Db    781  GATTCACCAGTGGAACACCTTGCTAAGTTGCATAACGGCAATTTTATTTGCTACAA    840
Qy    281  ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAlaAla    300
Db    841  CGCACGCCAGGAGTTGCCCGCAGCCGCGCGCCACCCCGTATTGGATTTGATCATGGCAGCG    900
Qy    301  LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu    320
Db    901  TTGACCCCCCATCCACCGCAAAAACAGGCCTATGGTGTGACATTACCCACTTCAGTACTG    960
Qy    321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuAlaGlnLeuGluLeuAsnTrp    340
Db    961  TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTCGAGCTCAACTGG   1020
Qy    341  ThrLeuProGlyGlnThrProAspAsnThrProProGlyLeuValPheGluGlnArgTrp    360
Db   1021  ACGCTTCCCGGTCAGCCCGGATAACACGCCGCCAGGTGGCGTGGTTGAACGCTGG   1080
Qy    361  ArgArgLeuSerAspAsnSerGlnIleTrpIleGluValSerLeuValPheGlnThrLeuGln    380
Db   1081  CGTCGGCTAAGCGATAACAGCCAGTGATTCAGGTTTCGCGTCGTTCCAGACTTTACAG   1140
Qy    381  GlnMetArgAspLysLysThrProLeuSerLeuAsnThrProProGlyLeuValLysLeuThr    400
Db   1141  CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCGGCCCGGAGAGGTGAAAACTGACC   1200
Qy    401  LeuAlaGlyLeuGluArgAspAsnAlaGlnAsnGlnAspGlyMetCysSerLeuLeuAlaGlyPheThrGln    420
Db   1201  CTGGCAGGATGTGAAGAGCGAAATGCCAGGGCATGTTGTTCGTTGGCAGGTTTTACGCAA   1260
Qy    421  IleValAsnGluAlaArgIleProAlaCysSerLeuLeuSerHisHisHisHisHis    440
Db   1261  ATCGTGAATGAAGCACGCATACCGGCGGCTGCAGTTTGAGATCTCATCACCATCACCATCAC   1320
```

```
RESULT 6
US-10-933-115-1
; Sequence 1, Application US/10933115
; Publication No. US20050281792A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
```

```
            APPLICANT:  Zyskind, Judith
            APPLICANT:  Wall, Daniel
            APPLICANT:  Trawick, John
            APPLICANT:  Carr, Grant
            APPLICANT:  Yamamoto, Robert
            APPLICANT:  Forsyth, R.
            APPLICANT:  Xu, H.
            TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
            FILE REFERENCE: ELITRA.034A
            CURRENT APPLICATION NUMBER: US/10/282,122A
            CURRENT FILING DATE:  2003-02-20
            PRIOR APPLICATION NUMBER: 60/191,078
            PRIOR FILING DATE:  2000-03-21
            PRIOR APPLICATION NUMBER: 60/206,848
            PRIOR FILING DATE:  2000-05-23
            PRIOR APPLICATION NUMBER: 60/207,727
            PRIOR FILING DATE:  2000-05-26
            PRIOR APPLICATION NUMBER: 60/230,335
            PRIOR FILING DATE:  2000-09-06
            PRIOR APPLICATION NUMBER: 60/230,347
            PRIOR FILING DATE:  2000-09-09
            PRIOR APPLICATION NUMBER: 60/242,578
            PRIOR FILING DATE:  2000-10-23
            PRIOR APPLICATION NUMBER: 60/253,625
            PRIOR FILING DATE:  2000-11-27
            PRIOR APPLICATION NUMBER: 60/257,931
            PRIOR FILING DATE:  2000-12-22
            PRIOR APPLICATION NUMBER: 60/267,636
            PRIOR FILING DATE:  2001-02-09
            PRIOR APPLICATION NUMBER: 60/269,308
            PRIOR FILING DATE:  2001-02-16
            Remaining Prior Application data removed - See File Wrapper or PALM.
            NUMBER OF SEQ ID NOS: 78614
            SOFTWARE: PatentIn version 3.1
            SEQ ID NO 7167
            LENGTH: 1299
            TYPE: DNA
            ORGANISM: Escherichia coli
            US-10-282-122A-7167

            Alignment Scores:
            Pred. No.:                              Length:         1299
            Score:                 1.12e-253        Matches:         430
            Percent Similarity:    2235.00          Conservative:      0
            Best Local Similarity: 99.5%            Mismatches:        2
            Query Match:           99.5%            Indels:            0
            DB:                    97.1%            Gaps:              0
                                   8

            US-10-601-319-2 (1-440) x US-10-282-122A-7167 (1-1299)

    Qy      1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla  20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db      1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA  60

    Qy     21 PheAlaGlnSerGluProGluLeuLysSerValValIleSerValValIleSerArgHisGly  40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db     61 TTCGCTCAGAGTGAGCCGGAGCTGAAGTCGGAAAGTGGTGAGATTGATTGTCAGTCGTCATGGT 120

    Qy     41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnArgValThrProAlaTrpPro  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db    121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGAGATGCACCCCAGACGCATGGCCA  180

    Qy     61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleLeuAlaTyrLeu  80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db    181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGCGGTGGTGAGCTAATCCGCCTATCTC 240

    Qy     81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro  100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db    241 GGACATTACCAACGCCACGCTCGGTAGCCGACCGGATTGTCGGCGAAAAGGGCTGCCCG  300

    Qy    101 GlnSerGlyGlnValAlaIleIleAlaAspValAlaAspGluAlaArgThrArgLysThrGlyGlu  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db    301 CAGTCTGGTCAGGTCGGCGATTATTGCTGATGTGACGAGCGTACCCGTAAAACAGGGCGAA  360
```

```
    Db    361 GCCTTCGCCGCCGGGCTGCACCTGACTGCTGCAATAACCGTACATACCCAGGCAGATACG  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Qy    141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysValGlnLeuAsnAspAsnAla  160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db    421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG  480

    Qy    161 AsnValThrAspAlaIleLeuSerArgAlaGlyAlaGlySerIleAlaGlyPheThrGlyHis  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db    481 AACGTGACTGACGCGATCCTCAGCAGGGCAGAGGGTCAATTGCTGACTTTACCGGGCAT  540

    Qy    181 ArgGlnThrAlaPheArgGluArgGlyLeuGlyLeuValLeuAsnPheProGlnSerArgLeuCys  200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db    541 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC  600

    Qy    201 LeuLysArgGlyLeuLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db    601 CTTAAACGTGGAGAAACAGGACGAAAGCGTGTTCATTAACGCAGGCATTACCATCGGAACTC  660

    Qy    221 LysValSerAlaAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db    661 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG  720

    Qy    241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrThrGlyArgIleThr  260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db    721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGTGGGAAGGATCACC  780

    Qy    261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln  280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db    781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA  840

    Qy    281 ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db    841 CGCACGCCAGAGGTTGCCCGCGCCAGCCGCCACCCCGTTATTGGATTTGATCATGGCAGCG  900

    Qy    301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu  320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db    901 TTGACGCCCCATCCACCGCAAAAACAGGCTATGGTGTGACATTACCCACTTCAGTACTG  960

    Qy    321 PheIleAlaGlyHisAspThrArgAsnLeuAlaAsnLeuGlyValAlaLeuGluLeuAsnTrp  340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db    961 TTTATTGCCGGACACGATACTCGTAATCTGCAAATCTCGCGGCGCCACTGGAGCTCAACTGG 1020

    Qy    341 ThrLeuProGlyGlnProAspAsnThrProProGlyProGlyLeuGlyLeuValPheGluArgTrp  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db   1021 ACGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGTGGTGAACTGGTGTTGAACGCTGG 1080

    Qy    361 ArgArgLeuSerAspAsnAspSerGlnAsnTrpIleGluSerLeuValPheGlnLnThrLeuGln  380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db   1081 CGTCGACTGGAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTGTCTTCCAGACTTTACAG 1140

    Qy    381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyGluValLysLeuThr  400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db   1141 CAGATGCGCTGATAAAACCCCGTGTCATTAAATAACGCGCCCGAGAGGTGAAACTGACC 1200

    Qy    401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db   1201 CTGGCAGGATGTGAAGAGCGAAATGCCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA 1260

    Qy    421 IleValAsnGluAlaArgIleProAlaCysSerSerLeuArgSerHisHisHisHisHis  440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db   1261 ATCGTTGAATGAAGCACGCATACCGGCGGTGCAGTTGAGATCTCATCACCACCATCAC 1320
```

```
RESULT 7
US-10-282-122A-7167
; Sequence 7167, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
   APPLICANT:  Wang, Liangbu
   APPLICANT:  Zamudio, Carlos
   APPLICANT:  Malone, Cheryl
   APPLICANT:  Haselbeck, Robert
   APPLICANT:  Ohlsen, Karl
```

```
;   APPLICANT:  GRAY, Kevin
;   APPLICANT:  BARTON, Nelson
;   APPLICANT:  GARRETT, James
;   APPLICANT:  O'DONOGHUE, Eileen
;   TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
;   FILE REFERENCE: DIVER1370-7
;   CURRENT APPLICATION NUMBER: US/09/866,379
;   CURRENT FILING DATE:  2001-05-24
;   PRIOR APPLICATION NUMBER: US 09/580,515
;   PRIOR FILING DATE: 2000-05-25
;   PRIOR APPLICATION NUMBER: US 09/318,528
;   PRIOR FILING DATE: 1999-05-25
;   PRIOR APPLICATION NUMBER: US 09/291,931
;   PRIOR FILING DATE: 1999-04-13
;   PRIOR APPLICATION NUMBER: US 09/259,214
;   PRIOR FILING DATE: 1999-03-01
;   PRIOR APPLICATION NUMBER: US 08/910,798
;   PRIOR FILING DATE: 1997-08-13
;   NUMBER OF SEQ ID NOS: 10
;   SOFTWARE: PatentIn version 3.1
;
;   SEQ ID NO 7
;   LENGTH: 1901
;   TYPE: DNA
;   ORGANISM: Escherichia coli
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1901)
;   OTHER INFORMATION: n is any nucleotide
US-09-866-379-7

Alignment Scores:
Pred. No.:                               1.95e-253      Length:                1901
Score:                                   2235.00        Matches:               430
Percent Similarity:                      99.5%          Conservative:          0
Best Local Similarity:                   99.5%          Mismatches:            2
Query Match:                             97.1%          Indels:                0
DB:                                             3        Gaps:                 0

US-10-601-319-2 (1-440) x US-09-866-379-7 (1-1901)

Qy      1  MetLysAlaLeuLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla        20
Db    188  ATGAAAGCGATCTTAATCCCATTTTATCTCTTCGTATCCGTTAACCCGCAATCTGCA         247

Qy     21  PheAlaGlnSerGluProGluLeuLysGluSerValValIleValSerArgHisGly         40
Db    248  TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT        307

Qy     41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro        60
Db    308  GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCGACGCATGGCCA         367

Qy     61  ThrTrpProValLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu         80
Db    368  ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGGNGTGGGTGAGCTATCGCCTATCTC       427

Qy     81  GlyHisTyrArgGlnProLeuAsnArgGlnAspCysAlaIleThrValHisThrGlnAlaAspThr        100
Db    428  GGACATTACCAACGCCAGCCGTCTGGATNGCCAGGATTGCTGGCGAAAAGGCTGCCCG         487

Qy    101  GlnSerGlyValGlnValAlaIleIleAlaAspValAspGlyArgThrLysThrGlyGlu        120
Db    488  CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACAGCCGTACCCGTAAAACAGGCCGAA       547

Qy    121  AlaPheAlaAlaGlyValAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr        140
Db    548  GCCTTCGCCGCCGGGCTGGCACCTGCAATAACCGTACATACCCAGGCAGATACG         607

Qy    141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla        160
Db    608  TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG       667

Qy    161  AsnValThrAspAlaArgAlaLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis        180
```

Qy   121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr   140
Db   361 GCCTTCGCCGCCGGGCTGGCACCTGCAATAACCGTACATACCCAGGCAGATACG   420

Qy   141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla   160
Db   421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG   480

Qy   161 AsnValThrAspAlaArgAlaLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis   180
Db   481 AACGTGACTGACGCGCGTGCCTTATCTCGAGCGGGTGGCAGTATTGCTGACTTTACCGGGCAT   540

Qy   181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys   200
Db   541 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC   600

Qy   201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu   220
Db   601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC   660

Qy   221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr   240
Db   661 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCCATCAATGCTGACG   720

Qy   241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr   260
Db   721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCCGAGCCGGGGTGGGGAAGGATCACC   780

Qy   261 AspSerHisGlnThrAsnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln   280
Db   781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAAGCGGTTGACATTAAGCCGGAATT   840

Qy   281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla   300
Db   841 CGGACACCGCCAGAGGTTGCCCGCAGCCCGCCACCCGTTATTAGATTTGATCGAAGACAGCG   900

Qy   301 LeuThrProHisProLysGlnIleLysGlnAlaTyrGlyValThrLeuProThrSerValLeu   320
Db   901 TTGACGCCCCATCCACCGCAAAAACAAGGCCGTATGGTGTGACATTACCCACTTCAGTGCTG   960

Qy   321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp   340
Db   961 TTTATCGCAGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG   1020

Qy   341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlyLeuValPheGluArgTrp   360
Db   1021 ACGCTTCCCGGTCAGCCCGGATAACACGCCGCCAGGTGGTGGTTTGAACCTGTTTGAACCCTGG   1080

Qy   361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln   380
Db   1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG   1140

Qy   381 GlnMetMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr   400
Db   1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCCGGAGAGGTGAAACTGACC   1200

Qy   401 LeuAlaGlyCysGluGluArgAsnAlaGlnIleGlyMetCysSerSerLeuAlaGlyPheThrGln   420
Db   1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGATTGGCATGTGTTCGTTCGTTGGCAGGTTTTACGCAA   1260

Qy   421 IleValAsnGluLeuAlaArgIleProAlaCysSerLeu   432
Db   1261 ATCGTGAATGAAGCACGCATACCGGCCTGCAGTTTG   1296

RESULT 8
US-09-866-379-7
; Sequence 7, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
;   APPLICANT: DIVERSA CORPORATION
;   APPLICANT: SHORT, Jay
;   APPLICANT: KRETZ, Keith

```
;   PRIOR FILING DATE: 2001-05-24
;   PRIOR APPLICATION NUMBER: US 09/580,515
;   PRIOR FILING DATE: 2000-05-25
;   PRIOR APPLICATION NUMBER: US 09/318,528
;   PRIOR FILING DATE: 1999-05-25
;   PRIOR APPLICATION NUMBER: US 09/291,931
;   PRIOR FILING DATE: 1999-04-13
;   PRIOR APPLICATION NUMBER: US 09/259,214
;   PRIOR FILING DATE: 1999-03-01
;   PRIOR APPLICATION NUMBER: US 08/910,798
;   PRIOR FILING DATE: 1997-08-13
;   NUMBER OF SEQ ID NOS: 4
;   SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 3
;   LENGTH: 1901
;   TYPE: DNA
;   ORGANISM: Escherichia coli
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (188)...(1483)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 403
;   OTHER INFORMATION: n = A,T,C or G
US-10-156-660-3
```

```
Alignment Scores:
Pred. No.:                                  Length:               1901
Score:                      1.95e-253       Matches:               430
Percent Similarity:         2235.00         Conservative:            0
Best Local Similarity:      99.5%           Mismatches:              2
Query Match:                99.5%           Indels:                  0
DB:                         97.1%           Gaps:                    0
US-10-156-660-3             6
```

US-10-601-319-2 (1-440) x US-10-156-660-3 (1-1901)

```
Qy      1  MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla  20
Db    188  ATGAAAGCGGATCTTAATCCCCATTTTTATCTTCTGATCCGTTAACCCGTAACTCTGCA  247

Qy     21  PheAlaGlnSerGluProGluLysLeuGluSerValValIleSerValArgHisGly  40
Db    248  TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGTGGGAAAGTGTGGATTGTCAGTCGTCATGGT  307

Qy     41  ValArgAlaProThrArgAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro  60
Db    308  GTGCGTGCTCCAACCAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGATGGCCA  367

Qy     61  ThrTrpProValLysGlyLeuGlyTrpLeuThrProArgGlyGluLeuIleAlaTyrLeu  80
Db    368  ACCTGGCCGGTAAAACTGGGTTGGCTGACACCCGCNGGTGGTGAGCTAATCGCCTATCTC  427

Qy     81  GlyHisTyrGlnArgGlnAlaIleIleIleAlaAspValAspGluThrArgLysThrGlyGlu  100
Db    428  GGACATTACCAACGCCAGGTCTTGGTAGCCGACGATTGCTCGGCGAAAAGGGCTGCCCG  487

Qy    101  GlnSerGlyGlnValAlaIleIleIleIleAlaAspValAspGluThrArgLysThrGlyGlu  120
Db    488  CAGTCTGGTCAGGTCGCCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA  547

Qy    121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaAlaIleThrValHisThrGlnAlaAspThr  140
Db    548  GCCTTCGCCGCCGGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG  607

Qy    141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla  160
Db    608  TCCAGTCCCGATCCGTTATTATCCCTAAAAACTGGCGTTTGCCAACTGGATAACGCG  667

Qy    161  AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaGlyPheThrGlyHis  180
Db    668  AACGTGACTGACGCGATCCTGACAGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT  727

Qy    181  ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys  200
```

Left column (continued alignment):

```
Db    668  AACGTGACTGACGCGATCCTCAGCAGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT  727

Qy    181  ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys  200
Db    728  CGGCAAACGGCGTTTCGGCAACTGGAACTGGAACGGGTGCTTAATTTCCCGCAATCAAACTTGTGC  787

Qy    201  LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnIleAlaLeuProSerGluLeu  220
Db    788  CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGATTGCTTTACCATCGGAACTC  847

Qy    221  LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr  240
Db    848  AAGGTGAGCGGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG  907

Qy    241  GluIlePheLeuLeuGlnGlnHisAlaGlnGlyMetProGluProGlyTyrTrpGlyTyrArgIleThr  260
Db    908  GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCCGGAGCCCGGGTGGGAGGGAGGATCACC  967

Qy    261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnAsnPheTyrLeuLeuGln  280
Db    968  GATTCACACCAGTGGAACACCCTGCTAAGTTTGCATAACGCCAATTTTATTTGCTACAA  1027

Qy    281  ArgThrProGluValAlaAlaSerArgAlaThrProLeuLeuAspLeuLeuIleMetAlaAlaAla  300
Db   1028  CGCACGCCAGAGGTTGCCGCAGCCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG  1087

Qy    301  LeuThrProHisProGlnLysGlnAlaTyrGlyValThrHisLeuProThrSerValLeu  320
Db   1088  TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCTG  1147

Qy    321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyValAlaGluLeuGluAsnTrp  340
Db   1148  TTTATCGCCGGACACGATACTAATCTCGGCAAATCTCGGCGCGGCCACTGGAGCTCAACTGG  1207

Qy    341  ThrLeuProGlyGlnProAspAsnThrProProGlyTyrGlyGluValPheGluLeuValArgTrp  360
Db   1208  ACGCTTCCCGGTCAGCCCGATAACACCGCCGGCCCAGGTGGTGAACTGGGTGTTGAACGGCTGG  1267

Qy    361  ArgArgLeuSerAspAsnSerGlnInTrpIleGlnValSerLeuValPheGlnInThrLeuGln  380
Db   1268  CGTCGGCTAAGCGATAACAGCCCAGTTGGATTCGAGTTTCGCTGGTCTTTCCAGACTTTTACAG  1327

Qy    381  GlnMetArgAspAspLysThrProLeuSerLeuAsnThrProProGlyTyrGluGluValLysLeuThr  400
Db   1328  CAGATGCGTGATAAAACGCCGCTGTCATTAAATACCGCCCGGAGAGGTGAAACTGACC  1387

Qy    401  LeuAlaGlyCysGluGluLeuArgAsnAlaGlnGlyAlaGlyLeuMetCysSerSerLeuAlaGlyPheThrGlyGln  420
Db   1388  CTGGCAGGATGTGAAGAGCGAAATGCCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA  1447

Qy    421  IleLeuValAsnGluAlaArgIleProAlaCysSerLeu  432
Db   1448  ATCGTGAATGAAGCACGGCATACCGGCCGCGTGCAGTTTG  1483
```

```
RESULT 9
US-10-156-660-3
;   Sequence 3, Application US/10156660
;   Publication No. US20030103958A1
;   GENERAL INFORMATION:
;   APPLICANT: Short, Jay M.
;   APPLICANT: Kretz, Keith
;   APPLICANT: Gray, Kevin A.
;   APPLICANT: Barton, Nelson R.
;   APPLICANT: Garrett, James B.
;   APPLICANT: O'Donoghue, Eileen
;   APPLICANT: Mathur, Eric J.
;   TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM
;   TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
;   FILE REFERENCE: 09010-029007
;   CURRENT APPLICATION NUMBER: US/10/156,660
;   CURRENT FILING DATE: 2002-10-01
;   PRIOR APPLICATION NUMBER: US 09/866,379
```

; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (188)...(1483)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 403
; OTHER INFORMATION: n = A,T,C or G
US-10-601-319-7

Alignment Scores:
Pred. No.:   1901
Score: 2235.00   Matches: 430
Percent Similarity: 99.5%   Conservative: 0
Best Local Similarity: 99.5%   Mismatches: 2
Query Match: 97.1%   Indels: 0
DB: 8   Gaps: 0

US-10-601-319-2 (1-440) x US-10-601-319-7 (1-1901)

| Qy | 1 | MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla | 20 |

| Db | 188 | ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATCCCGCTAACCCCGCAATCTGCA | 247 |

| Qy | 21 | PheAlaGlnSerGlyAspGluProLysLeuLysLeuGluSerValValIleValSerArgHisGly | 40 |

| Db | 248 | TTCGCTCAGAGTGGAGACGAGCCGAAGCTGGAAAAGTGTGGTGAGTCAGTCGTCATGGT | 307 |

| Qy | 41 | ValArgAlaProThrLysMetGlnAsnMetGlnAspValThrProAspAlaTrpProPro | 60 |

| Db | 308 | GTGCGTGCTCCAACCAAGATGCAGAACATGCAGGACGTCACCCCAGACGCATGCGCCA | 367 |

| Qy | 61 | ThrTrpProValLysGluGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu | 80 |

| Db | 368 | ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC | 427 |

| Qy | 81 | GlyHisTyrGlnArgGlnArgLeuValAlaAspValAlaAspGlyLysThrArgGlyThrGlyGlu | 100 |

| Db | 428 | GGACATTACCAACGCCAGCGTCTGGTAGCCGACGANGGCTGGCGGAAAAAGGGCTGCCCG | 487 |

| Qy | 101 | GlnSerGlyValAlaAlaIleIleAlaAspValAlaAspGluArgGlyThrGlyGluGlu | 120 |

| Db | 488 | CAGTCTGGTCAGGTCGCGGATCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACACAGGCGAA | 547 |

| Qy | 121 | AlaPheAlaAlaAlaGlyLeuAlaGlyAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr | 140 |

| Db | 548 | GCCTTCGCCGCCGGCGGGCCTGGCACCTGCAATAACCGTACATACCCAGGCAGATACG | 607 |

| Qy | 141 | SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla | 160 |

| Db | 608 | TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGTGTTTGCCAGCTGGATAACGCG | 667 |

| Qy | 161 | AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis | 180 |

| Db | 668 | AACGTGACTGACGCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTTACCGGGCAT | 727 |

| Qy | 181 | ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys | 200 |

| Db | 728 | CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC | 787 |

| Qy | 201 | LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu | 220 |

| Db | 728 | CGGCAAACGGCGTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC | 787 |

| Qy | 201 | LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu | 220 |

| Db | 788 | CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC | 847 |

| Qy | 221 | LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr | 240 |

| Db | 848 | AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCCGTAAGCCTCGCATCAATGCTGACG | 907 |

| Qy | 241 | GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr | 260 |

| Db | 908 | GAGATATTTCTCCTGCAACAGGCACAGGGAATGCCGGAGCCGGGGTGGGGAAGGATCACC | 967 |

| Qy | 261 | AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln | 280 |

| Db | 968 | GATTCACCAGTGGAACACCTTGCTAAGTTTGCATAACCGCAATTTTATTTGCTACAA | 1027 |

| Qy | 281 | ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla | 300 |

| Db | 1028 | CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTAGATTTGATCAAGAGACGCG | 1087 |

| Qy | 301 | LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu | 320 |

| Db | 1088 | TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCTG | 1147 |

| Qy | 321 | PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp | 340 |

| Db | 1148 | TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGGCGCACTGGAGCTCAACTGG | 1207 |

| Qy | 341 | ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp | 360 |

| Db | 1208 | ACGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG | 1267 |

| Qy | 361 | ArgArgLeuSerAspAsnSerGlnTrpIleIleGlnAsnValSerLeuValPheGlnThrLeuGln | 380 |

| Db | 1268 | CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGTCTTCCAGACTTTACAG | 1327 |

| Qy | 381 | GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyGluLeuValLeuThr | 400 |

| Db | 1328 | CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCGGAGAGGTGAAACTGACC | 1387 |

| Qy | 401 | LeuAlaGlyCysGluArgAspAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln | 420 |

| Db | 1388 | CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA | 1447 |

| Qy | 421 | IleValAsnGluAlaArgIleProAlaCysSerLeu | 432 |

| Db | 1448 | ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTG | 1483 |

RESULT 10
US-10-601-319-7
; Sequence 7, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING THEM
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528

```
;  PRIOR FILING DATE: 1999-03-01
;  PRIOR APPLICATION NUMBER: US 08/910,798
;  PRIOR FILING DATE: 1997-08-13
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 7
;    LENGTH: 1901
;    TYPE: DNA
;    ORGANISM: Escherichia coli
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (188)...(1483)
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: 403
;    OTHER INFORMATION: n = A,T,C or G
US-10-933-115-7

Alignment Scores:
Pred. No.:                      1901        Length:            1901
Score:                       2235.00        Matches:            430
Percent Similarity:            99.5%        Conservative:         0
Best Local Similarity:         99.5%        Mismatches:           2
Query Match:                   97.1%        Indels:               0
DB:                              10         Gaps:                 0

US-10-601-319-2 (1-440) x US-10-933-115-7 (1-1901)
```

```
Qy      1  MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla  20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188  ATGAAAGCGATCTTAATCCCATTTTATCCTCTGATTCCGTTAACGCCGCAATCTGCA  247

Qy     21  PheAlaGlnSerGluProGlyLeuLysGluSerValValIleValSerArgHisGly  40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248  TTCGCTCAGAGTGAGCCGGGAGCTGAAGCTGGTAGTGTGGTGATTGTCAGTCGTCATGT  307

Qy     41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308  GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA  367

Qy     61  ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu  80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    368  ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC  427

Qy     81  GlyHisTyrGlnArgGlnAsnGlyLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro  100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428  GGACATTACCAACGCCAGGGTCTGGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG  487

Qy    101  GlnSerGlyGlnValAlaIleIleAlaAlaAspValAspGluArgThrArgLysThrGlyGlu  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488  CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA  547

Qy    121  AlaPheAlaAlaGlyLeuLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr  140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548  GCCTTCGCGCCGGGCTGGCACCTGACTGCCAATAACCGTACATACCCGAGCAGATACG  607

Qy    141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla  160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608  TCCAGTCCGGATCCGTTATTTAATCCTCTAAAAACTGGCGTTGCCAACTGGATAACGCG  667

Qy    161  AsnValThrAspAlaIleLeuSerArgAlaGlyLysGlySerIleAlaAspPheThrGlyHis  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    668  AACGTGACTGACGCGATCCTCAGCAGGGCCAGGAGGGTCAATTGCTGACTTTACCGGGCAT  727

Qy    181  ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys  200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    728  CGGCAAACGGCGTTCGCGGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC  787

Qy    201  LeuLysArgGluLysAspAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    788  CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC  847

Qy    221  LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr  240
```

```
Db    788  CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCCGGAACTC  847
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    221  LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    848  AAGGTGAGCGGCCGACAATGTCTCATTAACCGTGCGGTAAGCCTCGCATCAATGCTGACG  907

Qy    241  GluIlePheLeuLeuGlnAsnGlnAlaGlyMetProGluProGlyTrpGlyArgIleThr  260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908  GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGTGGGGAAGGATCACC  967

Qy    261  AspSerHisGlnThrTrpAsnThrLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln  280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968  GATTCACACCAGTGGAACACCTTGCTAAGTTGCATAACGCGCAATTTTATTTGCTACAA  1027

Qy    281  ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028  CGCACGCCAGAGGTTGCCCGCAGCCGCCACCCCGTTATTAGATTTGATCAAGACAGCG  1087

Qy    301  LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu  320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088  TTGACGGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCTG  1147

Qy    321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148  TTTATCGCCGGACACGATACTAATCTCGGCGGCGCCACTGGAGCTCAACTGG  1207

Qy    341  ThrLeuProGlyGlyGlnProGlyAspAsnThrProProGlyGlyLeuValPheGluArgTrp  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208  ACGCTTCCCGGTCAGCCCGGATAACACGCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG  1267

Qy    361  ArgArgLeuSerArgAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln  380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268  CGTCGGCTAAGCGATAACAGCAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG  1327

Qy    381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyValLysLysLeuThr  400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328  CAGATGCGTGATAAAACGCCCCTGTCATTAAATACCGCCCCGGAGAGGTGAAACTGACC  1387

Qy    401  LeuAlaGlyCysGluGluValArgAsnAlaGlnGlyMetCysSerSerLeuAlaGlyPheThrGln  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388  CTGGCAGGATGTGAAGAGCGAAATGCCGCAGGGCATGTGTTCTTCGTTGGCAGGTTTTTACGCAA  1447

Qy    421  IleValAsnGluAlaArgIleProAlaCysSerLeu  432
              ||||||||||||||||||||||||||||||||||||
Db   1448  ATCGTGAATGAAGCACGCATACCGGGCGTGCAGTTTG  1483
```

```
RESULT 11
US-10-933-115-7
;  Sequence 7, Application US/10933115
;  Publication No. US20050281792A1
;  GENERAL INFORMATION:
;    APPLICANT: Short, Jay M.
;    APPLICANT: Kretz, Keith A.
;    APPLICANT: Gray, Kevin A.
;    APPLICANT: Barton, Nelson Robert
;    APPLICANT: Garrett, James B.
;    APPLICANT: O' Donoghue, Eileen
;    APPLICANT: Mathur, Eric J.
;    TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
;    TITLE OF INVENTION: THEREOF
;    FILE REFERENCE: 564462001822
;    CURRENT APPLICATION NUMBER: US/10/933,115
;    CURRENT FILING DATE: 2004-09-01
;    PRIOR APPLICATION NUMBER: US/09/866,379
;    PRIOR FILING DATE: 2001-05-24
;    PRIOR APPLICATION NUMBER: US 09/580,515
;    PRIOR FILING DATE: 2000-05-25
;    PRIOR APPLICATION NUMBER: US 09/318,528
;    PRIOR FILING DATE: 1999-05-25
;    PRIOR APPLICATION NUMBER: US 09/291,931
;    PRIOR FILING DATE: 1999-04-13
;    PRIOR APPLICATION NUMBER: US 09/259,214
```

; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli appA phytase
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (403)..(403)
; OTHER INFORMATION: n is any nucleotide
US-11-056-354-3

Alignment Scores:
Pred. No.:   1901
Score: 1.95e-253   Matches: 2235.00
Percent Similarity: 99.5%   Conservative: 0
Best Local Similarity: 99.5%   Mismatches: 2
Query Match: 97.1%   Indels: 0
DB: 14   Gaps: 0

US-10-601-319-2 (1-440) x US-11-056-354-3 (1-1901)

Qy   1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla   20
Db   188 ATGAAAGCGATCTTAATCCCATTTTATCTTCTGATTCCTTAACCCCGCAATCTGCA   247

Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly   40
Db   248 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT   307

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro   60
Db   308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA   367

Qy   61 ThrTrpProValLysGluGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu   80
Db   368 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGCGGTGAGCTAATCGCCTATCTC   427

Qy   81 GlyHisHisTyrGlnArgGlnAsnLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro   100
Db   428 GGACATTACCAACGCCAGAATCTGGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG   487

Qy   101 GlnSerGlyGlnValAlaIleAlaAspAspAspAlaAspAlaAspArgThrArgLysThrGlyGlu   120
Db   488 CAGTCTGGTCAGGTGGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAAACAGGCGAA   547

Qy   121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr   140
Db   548 GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG   607

Qy   141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla   160
Db   608 TCCAGTCCGGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG   667

Qy   161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis   180
Db   668 AACGTGACTGACGCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT   727

Qy   181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys   200
Db   728 CGGCAAACGGCGTTTCGGGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC   787

Qy   201 LeuLysArgGlnLysAspAlaGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu   220
Db   788 CTTAAACGTCAGAAAGACGGAACAGAAAGTGTTCATTAACGCAGCATTACCATCGGAACTC   847

Qy   221 LysValSerAlaAspAsnAlaValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr   240
Db   848 AAGGTGAGCGGCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG   907

Qy   241 GluIlePheLeuLeuGlnGlnGlnGlnGlyMetProGluProGlyTrpGlyArgIleThr   260
Db   908 GAGATATTTCTCCTGCAACAGCAGGGAATGCCCGAGCCCGGGTGGGGAAGGATCAACC   967

Db   848 AAGGTGAGCGGCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG   907

Qy   241 GluIlePheLeuLeuGlnGlnGlnGlyMetProGluProGlyTrpGlyArgIleThr   260
Db   908 GAGATATTTCTCCTGCAACAGCAGGGAATGCCCGAGCCCGGGTGGGGAAGGATCAACC   967

Qy   261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln   280
Db   968 GATTCCACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA   1027

Qy   281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla   300
Db   1028 CGCACGCCCAGAGGTTGCCGCAGCCGCGCCACCCCCGTTATTAGATTTGATCAAGACAGCG   1087

Qy   301 LeuThrProHisProProGlnLysGlyGlnAlaTyrGlyValThrLeuProThrSerValLeu   320
Db   1088 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCTG   1147

Qy   321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp   340
Db   1148 TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG   1207

Qy   341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp   360
Db   1208 ACGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGTGGTGAACTGGTGTTTGAACCGTGG   1267

Qy   361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln   380
Db   1268 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG   1327

Qy   381 GlnMetArgAspProLeuProLeuSerLeuAsnThrProProGlyGluValLysLeuThr   400
Db   1328 CAGATGCGTGATAAAACCGCCCGTGTCATTAAATACGGCGCCCGAGGAGGTGAAACTGACC   1387

Qy   401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln   420
Db   1388 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA   1447

Qy   421 IleValAlaAsnGluAlaArgIleProAlaCysSerLeu   432
Db   1448 ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTG   1483

RESULT 12
US-11-056-354-3
; Sequence 3, Application US/11056354
; Publication No. US20050246780A1
; GENERAL INFORMATION:
; APPLICANT: SHORT, Jay M.
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin A.
; APPLICANT: BARTON, Nelson R.
; APPLICANT: GARRETT, James B.
; APPLICANT: O'DONOGHUE, Eileen
; APPLICANT: MATHUR, Eric J.
; TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM AND METHODS
; TITLE OF INVENTION: FOR MAKING AND USING THEM
; FILE REFERENCE: 564462001803/D1370-9C1
; CURRENT APPLICATION NUMBER: US/11/056,354
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 10/156,660
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798

```
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: 403
;  OTHER INFORMATION: n = A,T,C or G
US-10-601-319-6

Alignment Scores:
Pred. No.:                  1.31e-252        Length:              1901
Score:                      2228.00          Matches:             429
Percent Similarity:         99.3%            Conservative:        0
Best Local Similarity:      99.3%            Mismatches:          3
Query Match:                96.8%            Indels:              0
DB:                                          Gaps:                0
                            8

US-10-601-319-2 (1-440)  x US-10-601-319-6  (1-1901)

US-10-601-319-2 (1-440)  x US-10-601-319-6  (1-1901)

Qy      1  MetLysAlaIleLeuIlePheLeuSerLeuLeuIleProLeuThrProGlnSerAla   20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188  ATGAAAGCGATCTTAATCCCATTTTATCTTCTGATTCCTGTTAACCCGCCAATCTGCA  247

Qy     21  PheAlaGlnSerGluProGluLeuLysGluSerValValIleValSerArgHisGly   40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248  TTCGCTCAGAGTGAGCCGGAGCTGAAAGCTGGAAAGTGTGGTGATTGTCGTCATGGT  307

Qy     41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro   60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308  GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA  367

Qy     61  ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu   80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    368  ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGGCGAACTAATCGCCTATCTC  427

Qy     81  GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro  100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428  GGACATTACTGGCGCCAGCGTCTGGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG  487

Qy    101  GlnSerGlyGlnValAlaIleLeuIleAlaAspValAspGluArgThrArgLysThrGlyGlu  120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488  CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA  547

Qy    121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr  140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548  GCCTTCGCCGCCGGCCTGGCACCTGACTGCGCAATAACCGTACATACCCAGGCAGATACG  607

Qy    141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla  160
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608  TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG  667

Qy    161  AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis  180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    668  AACGTGACTGACGCGATCCTCAGCAGGGCGAGGAGGTCAATTGCTGCATTCACCGGCCAT  727

Qy    181  ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys  200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    728  CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTCCCGCAATCAAACTTGTGC  787

Qy    201  LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    788  CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC  847

Qy    221  LysValSerAlaAspAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr  240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    848  AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTGGCCATCAATGCTGACG  907

Qy    241  GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr  260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908  GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCCGGGGTGGGAAGGATCACC  967

Qy    261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln  280
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968  GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA  1027

Qy    281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla  300
```

```
Db    908  GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCCGGGGTGGGAAGGATCACC  967

Qy    261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln  280
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968  GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA  1027

Qy    281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla  300
              ||||||||||||||||||||||||||||||||||||||||||         |||
Db   1028  CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTAGATTTGATCAAGAGACAGCG  1087

Qy    301  LeuThrProHisProProGlnLysGlyGlnAlaTyrGlyValThrLeuProThrSerValLeu  320
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088  TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCAGTCCAGTGCTG  1147

Qy    321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyValAlaLeuGluLeuAsnTrp  340
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148  TTTATCGCCGGACACGATACTAATCTGGCAAAATCTCGGCGGCGACGTGGAGCTCAACTGG  1207

Qy    341  ThrLeuProGlyGlnProAspAsnThrProGlyGlyGlyGluLeuValPheGluLeuArgTrp  360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208  ACGCTTCCCGGTCAGCCCGATAACACGCCCGGAGGTGGTGAACTGGTGTTTGAACGCTGG  1267

Qy    361  ArgArgLeuSerAspAsnGlnSerGlnIleGlnSerLeuValPheGlnThrLeuGln  380
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268  CGTCGGCTAAGCGATAACCAGAGCCAGTGATTCAGGTTCGTGTGTTTCCAGACTTTACAG  1327

Qy    381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValValLysLeuThr  400
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328  CAGATGCGTGATAAAAACGCCGTGTCATTAAATACGCCGCCCGGAGAGGTGAAACTGACC  1387

Qy    401  LeuAlaGlyCysGluGluArgAsnAlaGlnAlaGlyMetCysSerLeuAlaGlyPheThrGln  420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388  CTGGCAGGATGTGAAGAGCGAAATGCGCAGGCGGGCCATGTGTTCGTTGCAGGTTTTACGCAA  1447

Qy    421  IleValAsnGluAlaArgIleProAlaCysSerLeu  432
              |||||||||||||||||||||||||||||||||||||
Db   1448  ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTG  1483
```

```
RESULT 13
US-10-601-319-6
;  Sequence 6, Application US/10601319
;  Publication No. US20040091968A1
;  GENERAL INFORMATION:
;  APPLICANT: Short, Jay M.
;  APPLICANT: Kretz, Keith A.
;  APPLICANT: Gray, Kevin A.
;  APPLICANT: Barton, Nelson Robert
;  APPLICANT: Garrett, James B.
;  APPLICANT: O' Donoghue, Eileen
;  APPLICANT: Mathur, Eric J.
;  TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
;  TITLE OF INVENTION: AND USING THEM
;  FILE REFERENCE: 09010-029011
;  CURRENT APPLICATION NUMBER: US/10/601,319
;  CURRENT FILING DATE:  2003-06-20
;  PRIOR APPLICATION NUMBER: US 09/866,379
;  PRIOR FILING DATE: 2001-05-24
;  PRIOR APPLICATION NUMBER: US 09/580,515
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: US 09/318,528
;  PRIOR FILING DATE: 1999-05-25
;  PRIOR APPLICATION NUMBER: US 09/291,931
;  PRIOR FILING DATE: 1999-04-13
;  PRIOR APPLICATION NUMBER: US 09/259,214
;  PRIOR FILING DATE: 1999-03-01
;  PRIOR APPLICATION NUMBER: US 08/910,798
;  PRIOR FILING DATE: 1997-08-13
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 6
;  LENGTH: 1901
;  TYPE: DNA
;  ORGANISM: Escherichia coli
```

```
Best Local Similarity:    99.3%          Mismatches:    3
Query Match:              96.7%          Indels:        0
DB:                       3              Gaps:          0

US-10-601-319-2 (1-440) x US-09-866-379-9 (1-1901)

Qy      1 MetLysAlaIleLeuIleLeuProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla    20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188 ATGAAAGCGATCTTAATCCTCCATTTTATCTCTTCTGATTCCGTTAACCCGCCAATCTGCA      247

Qy     21 PheAlaGlnSerGluProGluLeuLysGluGlySerValValIleValSerArgHisGly       40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248 TTCGCTCAGAGTGAGCCGGAGCTGAAGGAAGGTGGTGATTGTCAGTCGTCATGT          307

Qy     41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro       60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA       367

Qy     61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu       80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    368 ACCTGGCCTGTAAAACTGGGTTGGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC       427

Qy     81 GlyHisTyrGlnArgGlnLeuAlaAspGlyLeuLeuAlaLysLysGlyCysPro            100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428 GGACATTACCAACGCCAGCTGTCGGTAGCCGACGGACTGCTGGCGAAAAAGGGCTGCCCG       487

Qy    101 GlnSerGlyGlnValAlaIleIleAlaAspValAspAspCysArgThrArgGlyLysGlu      120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA     547

Qy    121 AlaPheAlaAlaGlyLeuLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr   140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548 GCCTTCGCCGCCGGCGTGGCACTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG   607

Qy    141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla      160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608 TCCAGTCCGGATCCGTTATTTAATCCTCTAAAAACTGGCGTTGCCAACTGGATAACGCG       667

Qy    161 AsnValThrAspAlaAlaIleLeuSerArgAlaGlyGlySerIleIleAlaGlyHis         180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    668 AACGTGACTGACGCCGATCCTCAGCAGCGCAGGAGGTCAATTGCTGACTTTACCGGGCAT       727

Qy    181 ArgGlnThrAlaPheArgGluGluValGluValLeuAsnPheProGlnSerAsnLeuCys      200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    728 CGGCAAACGGCGTTTCGGAAGAGGTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC    787

Qy    201 LeuLysArgGluLysGlnAspArgSerCysSerLeuThrGlnAlaLeuProSerGluLeu      220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    788 CTTAAACGTGAGAAACAGGACAGAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC      847

Qy    221 LysValSerAlaAspAsnValSerLeuThrGlyValAlaValSerLeuAlaSerMetLeuThr   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    848 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCCATCAATGCTGACG     907

Qy    241 GluIlePheLeuGlnIleLysGlnIleAlaGlnMetProGlyLeuProGlyTrpGlyArgIleThr   260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908 GAGATATTTTCTCTGCAACAGCACAGGAATGCCCGGACCGGGCGGGGAGGATCACC         967

Qy    261 AspSerHisGlnTrpAsnThrLeuSerLeuHisSerAsnAlaGlnPheTyrLeuLeuGln      280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCCGCAATTTGTTTGCTCACAA     1027

Qy    281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuLeuMetAlaAla      300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028 CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG     1087

Qy    301 LeuThrProHisProGlnLysGlnGlnAlaTyrGlyValThrLeuProThrSerValLeu      320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGACATTACCCACTTCAGTGTGTCTG   1147

Qy    321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp      340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148 TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG     1207

Db   1028 CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG   1087

Qy    301 LeuThrProHisProGlnLysGlnGlnAlaTyrGlyValThrLeuProThrSerValLeu      320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGACATTACCCACTTCAGTGTGTCTG   1147

Qy    321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp      340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148 TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG   1207

Qy    341 ThrLeuProGlyGlnGlnAspThrAspProAspAsnThrProProGlyGlyGluLeuValPheTrp   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208 ACGCTTCCCGGTCAGCCGGATAACACGCCGCCGGATGGTGAACTGGTGTTTGAACGCTGG   1267

Qy    361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln      380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG   1327

Qy    381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysPheThrGln   400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328 CAGATGCGTGATAAAAACGCCCGTGTCATTAAATACGCCGCCGGAGAGGTGAAAACTGACC   1387

Qy    401 LeuAlaGlyCysGluGluArgAsnAlaGlnAsnGlyMetCysSerLeuAlaGlyPheThrGln   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA   1447

Qy    421 IleValAsnGluAlaArgIleProAlaCysSerLeu    432
          |||||||||||||||||||||||||||||||||||
Db   1448 ATCGTGAATGAATGAAGCACGCCATACCGGCCTGCAGTTTG   1483
```

RESULT 14
US-09-866-379-9
; Sequence 9, Application US/09866379
; Patent No. US2002013675A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-9

Alignment Scores:
Pred. No.:   2.27e-252   Length: 1901
Score: 2226.00   Matches: 429
Percent Similarity: 99.3%   Conservative: 0

```
Qy  101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  482 CAGCCTGGTCAGTTGCGGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA 541

Qy  121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  542 GCCTTCGCCGCCGGCTGGCACCTGCAATAACCGTACATACCCAGGCAGATACG 601

Qy  141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  602 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAACTGGGCGTTTGCCAACTGGATAACGCG 661

Qy  161 AsnValThrAspAlaIleLeuSerArgAlaAlaGlyServeIleAlaAspPheThrGlyHis 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  662 AACGTGACTGACGCGATCCTCAGCAGGGCAGGGGTCAATTGCTGACTTTACCGGGCAT 721

Qy  181 ArgGlnThrAlaPheArgGluArgGlnLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  722 CGGCAAACGCGTTTCGCAGGAACTGGAACGGGTGCTTAATTTTCCCAATAAACTTGTGC 781

Qy  201 LeuLysArgGluLysGlnAspGluSerCysSerSerLeuThrGlnAlaLeuProSerGluLeu 220
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  782 CTTAACCGTGAGAAACAGGACGAAAGCTGTTCATTAACCGGCGATTACCATCGGAACTC 841

Qy  221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerAlaSerMetLeuThr 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  842 AAGGTGAGCGCCGACAATGTTCATTAACCGGTGCGGTAAGCCTGCCATCAATGCTGACG 901

Qy  241 GluIlePheLeuLeuGlnGlnAlaGlnGlnMetProGluProGlyTrpGlyArgIleThr 260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  902 GAAATATTTCTCCTGCAACAAGCACAGCAGATGCCGGAGCCGGGGTGGGGAAGGATCACT 961

Qy  261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  962 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTACTACAA 1021

Qy  281 ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1022 CGCACGCCAGAGGTTGCCGCAGTCGCGGCCACCCCGTTATTGGATTTGATCATGGCAGCG 1081

Qy  301 LeuThrProHisProGlnAsnGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1082 TTGACGCCCCATCCACCGCAAAAACAGGGCTATGGTGTGACATTACCCACTTCAGTGCTG 1141

Qy  321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaAlaGluGlyLeuAsnTrp 340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1142 TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGGCGCCACTCGGAGCTCAACTGG 1201

Qy  341 ThrLeuProGlyGlnProAspAsnThrProGlyGlyValPheGluValPheGlnThrGln 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1202 ACGCTTCCAGGTCAGCCCGATAACACGCCCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1261

Qy  361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnThrLeuValSerLeuValPheGlnThrLeuGln 380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1262 CGTCGGCTAAGCGATAACAGCCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1321

Qy  381 GlnMetArgAspThrProLeuSerLeuAsnThrProProGlyValLysLeuThr 400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1322 CAGATGCGTGATAAAACGCCGCTATCATTAAAATACGCCGCCGGTGAAACTGACC 1381

Qy  401 LeuAlaGlyCysGluGluArgAsnAlaGlnLeuMetCysSerLeuAlaGlyPheThrGln 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1382 CTGGCAGGATGTGAAGAGCGAAATGCCAGGCGCATGTGTTCGTTGGCCGGTTTTACGCAA 1441

Qy  421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
            |||||||||||||||||||||||||||||||||||||
Db  1442 ATCGTGAATGAAGGCGCATACCGGCGTGCAGTTTG 1477
```

Search completed: June 14, 2006, 15:10:07
Job time : 1468.58 secs

```
Qy  341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1208 ACGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1267

Qy  361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnThrLeuValSerLeuValPheGlnThrGln 380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1268 CGTCGGCTAAGCGACAAACAGCCCGGATTGGAATTCCGTTTCGCTGGTCTTCCAGACTTTACAG 1327

Qy  381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysPheGluLeuThr 400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1328 CAGATGCGGTGATAAAAACGCCGCTGTCATTAAATACGCCGCCCGGAGAGGTGAAACTGACC 1387

Qy  401 LeuAlaGlyCysGluGluArgAsnAlaGlnLeuMetCysSerLeuAlaGlyPheThrGln 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1388 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGCGCATGTGTTCGTTGGCCAGGTTTTACGCAA 1447

Qy  421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
            ||||||||||||||||||||||||||||||||||||
Db  1448 ATCGTGAATGAAGAAGCACGCATACCGGCGTGCAGTTTG 1483
```

RESULT 15
US-10-266-041-9
; Sequence 9, Application US/10266041
; Publication No. US2003072844A1
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2791
; CURRENT APPLICATION NUMBER: US/10/266,041
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: US/09/540,149
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 60/127,032
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1489
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-266-041-9

Alignment Scores:
Pred. No.:   Length: 1489
Score: 6.22e-252   Matches: 428
Percent Similarity: 2221.00   Conservative: 0
Best Local Similarity: 99.1%   Mismatches: 4
Query Match: 99.1%   Indels: 0
DB: 96.5%   Gaps: 0
                6

US-10-601-319-2 (1-440) x US-10-266-041-9 (1-1489)

```
Qy    1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  182 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTTTGATTCCGTTAACCCCGCAATCTGCA 241

Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  242 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTAATTGTCAGCCGTCATGGT 301

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  302 GTGCGTGCGCCAACCAAGGCCACGCAACTGATGCAGGACGTGACCCCAGACGCATGGCCA 361

Qy   61 ThrTrpProValLysGlyLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  362 ACCTGGCCGGTAAACTGGGTTGGCTGACACCACGCGGTGGTGAGCTAATCGCCTATCTC 421

Qy   81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  422 GGACATTACCAACGCCAGCGTCTGGTGGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG 481
```

THIS PAGE LEFT BLANK

```
              GenCore version 5.1.9
        Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:      June 14, 2006, 12:41:52 ; Search time 224.541 Seconds
                                       (without alignments)
                                       5499.798 Million cell updates/sec

Title:           US-10-601-319-2
Perfect score:   2302
Sequence:        1 MKAILIPFISLLIPLTPQSA.........IVNEARIPACSLRSHHHHHH 440

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext 0.5
                 Ygapop 10.0 , Ygapext 0.5
                 Fgapop  6.0 , Fgapext 7.0
                 Delop   6.0 , Delext  7.0

Searched:        1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters:    2807332

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1p
-Q=/abss/ABSSWEB spool/US10601319/runat 12062006 180055 27578/runat 12062006 180055 27578/app_query.fasta_1
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss03p -USER=US10601319 @CGN_1_1_381 @runat_12062006_180055_27578
-NCPU=6 -ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  Issued Patents NA:*
            1: /EMC_Celerra_SIDS3/ptodata/2/ina/1_COMB.seq:*
            2: /EMC_Celerra_SIDS3/ptodata/2/ina/5_COMB.seq:*
            3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A_COMB.seq:*
            4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B_COMB.seq:*
            5: /EMC_Celerra_SIDS3/ptodata/2/ina/7_COMB.seq:*
            6: /EMC_Celerra_SIDS3/ptodata/2/ina/H_COMB.seq:*
            7: /EMC_Celerra_SIDS3/ptodata/2/ina/PCTUS_COMB.seq:*
            8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP_COMB.seq:*
            9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE_COMB.seq:*
           10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfiles1.seq:*
```

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2302	100.0	1323	3	US-09-259-214-1	Sequence 1, Appli
2	2302	100.0	1323	3	US-09-318-528-1	Sequence 1, Appli
3	2302	100.0	1323	3	US-09-291-931-1	Sequence 1, Appli
4	2302	100.0	1323	3	US-09-580-515-1	Sequence 1, Appli
5	2302	100.0	1323	3	US-09-866-379-1	Sequence 1, Appli
6	2235	97.1	1323	3	US-09-866-379-7	Sequence 7, Appli
7	2226	96.7	1901	3	US-09-866-379-9	Sequence 9, Appli
8	2221	96.5	1489	3	US-09-540-149A-9	Sequence 9, Appli
9	2221	96.5	1489	4	US-10-266-041A-9	Sequence 9, Appli
10	2221	96.5	1901	3	US-09-866-379-6	Sequence 6, Appli
11	2220	96.4	1901	3	US-09-866-379-5	Sequence 5, Appli
12	2217	96.3	1486	3	US-09-715-477-2	Sequence 2, Appli
13	2197	95.4	1486	2	US-09-715-477-4	Sequence 4, Appli
14	2194.5	95.3	1272	2	US-08-910-798-1	Sequence 1, Appli
15	528	22.9	1266	3	US-09-489-039A-341	Sequence 341, App
16	376	16.3	1569	3	US-09-489-039A-6330	Sequence 6330, App
17	352	15.3	5975	2	US-09-920-812-23	Sequence 23, Appl
18	352	15.3	5975	2	US-08-920-827-23	Sequence 23, Appl
19	352	15.3	5975	3	US-08-921-177-23	Sequence 23, Appl
20	352	15.3	5975	2	US-08-362-577C-23	Sequence 23, Appl
21	352	15.3	5975	2	US-08-920-828-23	Sequence 23, Appl
22	119	5.2	1571	3	US-08-868-435-32	Sequence 32, Appl
23	119	5.2	1571	3	US-08-744-231-32	Sequence 32, Appl
24	119	5.2	1571	3	US-09-044-718-10	Sequence 10, Appl
25	119	5.2	1571	3	US-09-635-504-32	Sequence 32, Appl
26	119	5.2	1571	3	US-10-062-848-10	Sequence 10, Appl
27	117	5.1	1588	2	US-09-146-283-1	Sequence 1, Appli
28	117	5.1	1588	3	US-08-579-823A-1	Sequence 1, Appli
29	117	5.1	1588	3	US-09-344-195-1	Sequence 1, Appli
30	117	5.1	2283	3	US-09-252-991A-2140	Sequence 2140, Ap
31	117	5.1	3061	2	US-08-692-787-47	Sequence 47, Appl
32	117	5.1	3061	3	US-09-097-199-47	Sequence 47, Appl
33	117	5.1	3089	3	US-09-949-016-151	Sequence 151, App
34	115.5	5.0	1336	3	US-09-949-016-3480	Sequence 3480, Ap
35	115.5	5.0	1336	3	US-09-949-016-3481	Sequence 3481, Ap
36	115.5	5.0	1336	3	US-09-949-016-3482	Sequence 3482, Ap
37	115.5	5.0	1336	3	US-09-949-016-3483	Sequence 3483, Ap
38	115.5	5.0	1336	3	US-09-949-016-3484	Sequence 3484, Ap
39	115.5	5.0	1372	3	US-09-765-111A-13	Sequence 13, Appl
40	115	5.0	7404	3	US-09-902-540-3115	Sequence 3115, Ap
41	115	5.0	18809	3	US-09-902-540-1141	Sequence 1141, Ap
42	114.5	5.0	1758	5	US-10-144-678A-1019	Sequence 1019, Ap
43	113.5	4.9	2625	3	US-09-765-111A-5	Sequence 5, Appli
44	113.5	4.9	2711	3	US-09-765-111A-24	Sequence 24, Appl
45	112	4.9	11604	3	US-09-385-028-13	Sequence 13, Appl

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

```
RESULT 1
US-09-259-214-1
; Sequence 1, Application US/09259214A
; Patent No. 6110719
; GENERAL INFORMATION:
;   APPLICANT: Kretz, Keith
;   TITLE OF INVENTION: NOVEL PHYTASE
;   FILE REFERENCE: DIVER1370-1
;   CURRENT APPLICATION NUMBER: US/09/259,214A
;   CURRENT FILING DATE: 1999-03-01
;   EARLIER APPLICATION NUMBER: 08/910,798
;   EARLIER FILING DATE: 1997-08-13
;   NUMBER OF SEQ ID NOS: 4
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1323
;   TYPE: DNA
;   ORGANISM: Escherichia coli
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(1320)
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1323)
;   OTHER INFORMATION: n = A,T,C or G
US-09-259-214-1

Alignment Scores:
Pred. No.:                  1.05e-233        Length:         1323
Score:                      2302.00          Matches:         440
Percent Similarity:         100.0%           Conservative:      0
Best Local Similarity:      100.0%           Mismatches:        0
Query Match:                100.0%           Indels:            0
```

```
DB:          3          Gaps:     0

US-10-601-319-2 (1-440) x US-09-259-214-1 (1-1323)

Qy    1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla  20
Db    1 ATGAAAGCGATCTTAATCCATTTTATCTCTCTGATTCCGTTAACCCCGCAATCTGCA  60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   21 PheAlaGlnSerGluProGluLeuLysGluSerValValIleValSerArgHisGly  40
Db   61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGGTGATTGTCAGTCGTCATGGT  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   41 ValArgAlaProThrLysMetGlnAspValThrProAspAlaTrpPro  60
Db  121 GTGCGTGCTCCAACCAAGATGCAGGATGTCACCCCAGACGCATGGCCA  180
        ||||||||||||||||||||||||||||||||||||||||||||||||

Qy   61 ThrTrpProValLysGluGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu  80
Db  181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGTGAGCTAATCGCCTATCTC  240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   81 GlyHisTyrGlnAsnArgGlnArgLeuValAlaAlaAspGlnLeuLeuAlaLysGlyCysPro  100
Db  241 GGACATTACCAACGCCAGCGTCTGGTAGCCGGCGGATTGCTGGCGAAAAGGGCTGCCCCG  300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  101 GlnSerGlyGlnValAlaIleIleAlaIleAspAspValAspGluArgThrArgLysThrGlyGlu  120
Db  301 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAAACAGGCGAA  360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  121 AlaPheAlaAlaGlyLysLeuAlaAlaProAspCysTyrAlaIleThrValHisThrGlnAlaAspThr  140
Db  361 GCCTTCGCCGCCGGGCTGGCCACCTGTGCAATACCGTGACATACCCAGGCAGATACG  420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  141 SerSerProAspAspPheAsnProLeuLysThrGlyTyrValCysGlnLeuAspAspAsnAla  160
Db  421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTGCCAACTGGATAACGCGG  480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  161 AsnValThrAspAlaPheArgAlaGlyLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis  180
Db  481 AACGTGACTGACGCGGATCCTCAGCAGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT  540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys  200
Db  541 CGGCAAACGGCGTTTCGGGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTGTGC  600
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  201 LeuLysArgGluArgLysGlnAspAspLeuSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220
Db  601 CTTAAACGTGAGAGAACAGGACGAAAGCTGTTCATTAACGCATTACCATCGGAACTC  660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValAlaValSerLeuAlaSerMetLeuThr  240
Db  661 AAGGTGAGCGCCGACAAATGTCTCATTAACGGTGCCGTAAGCCTCGCATCAATGCTGACG  720
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  241 GluIleGluPheLeuGlnLeuGlnLeuAsnGlyMetProGluProGlyTrpGlyArgIleThr  260
Db  721 GAAATCGAGTTCCTGCAACTGCAACTGAATGGGATGCCGGAGCCGGGTGGGAAGGATCACC  780
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  261 AspSerHisGlnThrPheAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln  280
Db  781 GATTCACACCAGTGACGAACACCTTGCTAAGTTGCATAACGCGCAATTTTATTTGCTACAA  840
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  281 GluIlePheGluLeuGlnLeuGlyGlnGlnAlaThrProLeuLeuAspLeuIleMetAlaAla  300
Db  841 CGCACGCCAGAGGTTGCCCGCAGCCGCCGCCACCCGTTATTGGATTTGATCATGGCAGCG  900
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  301 LeuThrProHisProProGlnLysGlnAlaThrTyrGlyValThrLeuProThrSerValLeu  320
Db  901 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGACATTACCCCACTCAGTACTG  960
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuGlnTrp  340
Db  961 TTTATTGCCGGACACGACACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG  1020
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlyLeuValPheGluLeuArgTrp  360
Db 1021 ACGCTTCCGGTCAGCCGGATAACACCGCCCAGGTGGTGAACTGGTGTTGAACGCTGG  1080
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  361 ArgArgLeuSerAspAsnSerGlnTrpIleGlyValSerLeuValPheGlnThrLeuGln  380
Db 1081 CGTCGGCTAAGGCGGATAACAGCCAGTGGATTGGATGATTCGGTCGTCTTACAG  1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProGlyProGlyValLysLeuThr  400
Db 1141 CAGATGCGTGATAAAACGCCGTCATTAAATACGGCCGCCCGGAGAGGTGAAACTGACC  1200
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln  420
Db 1201 CTGGCAGGATGTGAAGAGCGAAATGCCAGGGCATGTTCGTTGGCAGGTTTTACGCAA  1260
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  421 IleValAlaAsnGluAlaAsnArgIleProAlaCysSerLeuArgSerHisHisHisHisHis  440
Db 1261 ATCGTGAATGAAGCAGCACGCATACCGGCCTGCAGTTTGAGAATCTCATCACCACCATCAC  1320
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Alignment Scores:
Pred. No.:                                  Length:           1323
Score:                      1.05e-233       Matches:           440
                            2302.00         Conservative:        0
Percent Similarity:         100.0%          Mismatches:          0
Best Local Similarity:      100.0%          Indels:              0
Query Match:                100.0%          Gaps:                3
DB:

US-10-601-319-2 (1-440) x US-09-318-528-1 (1-1323)

Qy    1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla  20
Db    1 ATGAAAGCGATCTTAATCCATTTTATCTCTCTGATTCCGTTAACCCCGCAATCTGCA  60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   21 PheAlaGlnSerGluProGluLeuLysGluSerValValIleValSerArgHisGly  40
Db   61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGGTGATTGTCAGTCGTCGTCATGGT  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro  60
Db  121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA  180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   61 ThrTrpProValLeuGluGlyTrpThrProArgGlyLeuLeuThrProArgGlyLeuTyrPro  80
Db  181 GTGCGGTGTCTGGGTCTGGGTACTGCTTTTTATCTTCTTGATTCCGTTAACCCCGCAATCTGCA  240
```

RESULT 3
US-09-291-931-1
; Sequence 1, Application US/09291931A
; Patent No. 6190897
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/291,931A
; CURRENT FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; NAME/KEY: misc_feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-291-931-1

Alignment Scores:
Pred. No.:   Length: 1323
Score: 1.05e-233   Matches: 440
Percent Similarity: 2302.00   Conservative: 0
Best Local Similarity: 100.0%   Mismatches: 0
Query Match: 100.0%   Indels: 0
DB: 3   Gaps: 0

US-10-601-319-2 (1-440) x US-09-291-931-1 (1-1323)

Qy	1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuLeuLeuThrProGlnSerAla	20
Db	1 ATGAAAGCGATCTTAATCCCATTTTATCTCTTCTGATTCCTTAACCCCGCAATCTGCA	60
Qy	21 PheAlaGlnSerGluProGluLysLysGluSerValValIleValSerArgHisGly	40
Db	61 TTCCCTCAGAGTGAGCCGGAGCTGGAAGTGGTCGGTGATTGTCGTCGTCATGGT	120
Qy	41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro	60
Db	121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGATGGCCA	180
Qy	61 ThrTrpProValLysTrpLeuGlyTrpProGlnAlaLeuAlaTyrLeu	80
Db	181 ACCTGGCCGGTAAAACTGGGTTGGCCGCAGGCGCTGGCCTATCTC	240
Qy	81 GlyHisTyrGlnArgGlnArgLeuValAlaAspValArgGluLeuAlaLysGlyCysPro	100
Db	241 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGTGCGGGAAAAGGGCTGCCCG	300
Qy	101 GlnSerGlyGlnValAlaIleIleIleAlaAspThrAspArgThrArgArgThrGlyGlu	120
Db	301 CAGTCTGGTCAGTGCGGCGAATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA	360
Qy	121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr	140
Db	361 GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG	420
Qy	141 SerSerProAspProLeuPheAsnProLeuLysLysGlyValCysGlnLeuAspAsnAla	160
Db	421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG	480
Qy	161 AsnValThrAspAlaAlaLeuSerArgAlaGlyLySerIleLeuAlaAspPheThrGlyHis	180
Db	481 AACGTGACTGACGCCGATCCTCAGCAGGGCGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT	540

(left column, Db/Qy strands of query SEQ)

Db	181 ACCTGGGCCGGTAAAACTGGGTTGGCTGACACCCGCGGNGTGGTGAGCTAATCGCCTATCTC	240
Qy	81 GlyHisTyrGlnArgGlnArgLeuValAlaAspValArgGluLeuAlaLysGlyCysPro	100
Db	241 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGTGCGGGAAAAAGGGCTGCCCG	300
Qy	101 GlnSerGlyGlnValAlaIleIleIleAlaAspThrAspArgThrArgArgThrGlyGlu	120
Db	301 CAGTCTGGTCAGTGCGGCGAATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA	360
Qy	121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr	140
Db	361 GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG	420
Qy	141 SerSerProAspProLeuPheAsnProLeuLysLysGlyValCysGlnLeuAspAsnAla	160
Db	421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG	480
Qy	161 AsnValThrAspAlaAlaLeuSerArgAlaGlyLySerIleLeuAlaAspPheThrGlyHis	180
Db	481 AACGTGACTGACGCCGATCCTCAGCAGGGCGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT	540
Qy	181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys	200
Db	541 CGGCAAACGGCCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC	600
Qy	201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu	220
Db	601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGGAACTC	660
Qy	221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr	240
Db	661 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTGGCCATCAATGCTGACG	720
Qy	241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGlyLeuProGlyTyrArgIleThr	260
Db	721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGGCCGGGGTGGGAAGGATCACC	780
Qy	261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln	280
Db	781 GATTCCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGGCGCAATTTTATTGCTACAA	840
Qy	281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla	300
Db	841 CGCACCGCCAGAGGTTGCCCGCAGCCGGCGCACCCCCGTTATTGGATTTGATCATGGCAGCG	900
Qy	301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu	320
Db	901 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTACTG	960
Qy	321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp	340
Db	961 TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG	1020
Qy	341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp	360
Db	1021 ACGCTTCCCGGTCAGCCTGACAATACACCGCCAGGTGGTGAACTGGTGTTTGAACGGTGG	1080
Qy	361 ArgArgLeuSerAspAsnSerGlnInTrpIleGlnInThrGlnInThrLeuGln	380
Db	1081 CGTCGGCTAAGCGATAACAGCCAGTTCATTCAGATTCAGGTTTCGCTCGTCTTCCAGACTTTACAG	1140
Qy	381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLeuLeuThr	400
Db	1141 CAGATGCGCGATAAAACGCCGCTGTCATTAAATACGCCGCCCGGAGAGGTGAAACTGACC	1200
Qy	401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln	420
Db	1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCGTTGGCAGGTTTACGCAA	1260
Qy	421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHisHis	440
Db	1261 ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTGAGATCTCATCATCACCACCATCAC	1320

```
;     SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;     LENGTH: 1323
;     TYPE: DNA
;     ORGANISM: Escherichia coli
;     FEATURE:
;     NAME/KEY: CDS
;     LOCATION: (1)...(1320)
;     NAME/KEY: misc_feature
;     LOCATION: (1)...(1323)
;     OTHER INFORMATION: n = A,T,C or G
US-09-580-515-1

Alignment Scores:
Pred. No.:             1.05e-233          Length:        1323
Score:                 2302.00            Matches:        440
Percent Similarity:    100.0%             Conservative:     0
Best Local Similarity: 100.0%             Mismatches:       0
Query Match:           100.0%             Indels:           0
DB:                                       Gaps:             3

US-10-601-319-2 (1-440) x US-09-580-515-1 (1-1323)

Qy    1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla   20
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAGCGATCTTAATCCCATTTTTATCCTTGTTGATTCCGTTAACCCCGCAATCTGCA   60

Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly   40
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT  120

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro   60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA  180

Qy   61 ThrTrpProValLysTrpLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC  240

Qy   81 GlyHisTyrGlnArgGlnArgLeuValAlaAspAlaLysLeuLeuAlaLysGlyCysProPro 100
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGCGATTCTGCGGCGAAAAAGGCCTGCCCG 300

Qy  101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu  120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA  360

Qy  121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr  140
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GCCTTCGCCGCGGGCCTGGCACCTGACTGCGCAATAACCGTACATACCCAGGCAGATACG  420

Qy  141 SerSerProAspProLeuLysProLeuPheThrValCysGlnLeuAspAsnAla       160
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TCCAGTCCGGATCCGTTATTTAATCCTCTAAAAACTCGGCGTTTGCCAACTGGATAACGCG 480

Qy  161 AsnValThrAspAlaAspProGlnGlnAlaGlyGlySerIleAlaAspPheThrGlyHis  180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 AACGTGACTGACGCCGATCCTCAGCAGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT  540

Qy  181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys  200
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 CGGCAAACGGCGTTCCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC  600

Qy  201 LeuLysArgGluLysGlnAspSerCysSerLeuThrGlnIleAlaLeuProSerGluLeu  220
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGATTGCATTACCATCGGAACTC 660

Qy  221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr  240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG  720

Qy  241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluThrGlyLeuTrpGlyArgIleThr 260
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAAACGGGTGGGGAAGGAATCACC  780

Qy  261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln  280
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA  840

Qy  281 ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuAspLeuIleMetAlaAla  300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 CGCACGCCGGAGGTTGCCGCGCGCAGCCGCGCCACCCCGTTATTGGATTTGATCATGGCAGCG 900

Qy  301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProGluThrSerValLeu 320
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TTGACGCCCCATCCACCGCAAAAACAGGCCTATGGTGTGACATTACCCGAGACTTCAGTACTG 960

Qy  321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG 1020

Qy  341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp  360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 ACGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1080

Qy  361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140

Qy  381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr  400
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCGGGAGAGGTGAAACTGACC 1200

Qy  401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln  420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA 1260

Qy  421 IleValAsnGluAlaArgIleProAlaLysCysSerLeuArgSerHisHisHisHisHis  440
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 ATCGTGAATGAAGCACACCATACCGGCGCCTGCAGTTTGAGATCTCATCACCATCACCAC 1320
```

```
;   OTHER INFORMATION: n is any nucleotide
;   NAME/KEY: CDS
;   LOCATION: (1)..(1323)
;   OTHER INFORMATION:
US-09-866-379-1

Alignment Scores:
Pred. No.:                    1.05e-233      Length:           1323
Score:                        2302.00        Matches:          440
Percent Similarity:           100.0%         Conservative:     0
Best Local Similarity:        100.0%         Mismatches:       0
Query Match:                  100.0%         Indels:           0
DB:                           3              Gaps:             0

US-10-601-319-2 (1-440) x US-09-866-379-1 (1-1323)

Qy    1  MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuIleProGlnSerAla   20
Db    1  ATGAAAGCGATCTTAATCCCATTTTATCTTCTGATTCCGTTAACCCCGCAATCTGCA       60

Qy   21  PheAlaGlnSerGluProGluLeuLysGluSerValValIleValSerArgHisGly      40
Db   61  TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGTGGTGATTGTCAGTCGTCATGGT          120

Qy   41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro   60
Db  121  GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCGGACGCATGGCCA   180

Qy   61  ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyLeuLeuIleAlaTyrLeu   80
Db  181  ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGGACTAATCGCCTATCTC     240

Qy   81  GlyHisTyrGlnArgGlnArgLeuValLeuAlaAspGlyLeuLeuIleAlaLysGlyCysPro  100
Db  241  GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGCTGGGGAAAAAGGGGTGCCCG   300

Qy  101  GlnSerGlyGlnValAlaIleIleIleAlaAlaAspValAspGluArgThrArgLysThrGlyGlu  120
Db  301  CAGTCTGGTCAGGTCGCAGGTATTGTGATGTGCACGACGAGCGTACCCGTAAAACAGGCGA  360

Qy  121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr  140
Db  361  GCCTTCGCCGCCGGCTGGCACCTGCACATACCTACATAACCCAGCAGCAGATACG       420

Qy  141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla  160
Db  421  TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTGCCAACTGGATAACGCG    480

Qy  161  AsnValThrAspAlaIleLeuSerArgAlaGlyLeuGlySerIleAlaAspPheThrGlyHis  180
Db  481  AACGTGACTGACGCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT   540

Qy  181  ArgGlnThrAlaPheArgGluLeuGlnArgValLeuAsnPheProGlnSerAsnLeuCys  200
Db  541  CGGCAAACGGCGTTTCGCGAACTGGAACCGGGTGCTTAATTTTCCGCAATCAAACTTGTGC  600

Qy  201  LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220
Db  601  CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC   660

Qy  221  LysValSerAlaAspAspAsnValSerLeuThrGlyValAlaValSerAlaSerMetLeuThr  240
Db  661  AAGGTAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTGACGTAAGTGCTGACG   720

Qy  241  GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrArgIleThr  260
Db  721  GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGTGGGAAGGATCACC   780

Qy  261  AspSerHisGlnTrpAsnThrLeuAsnSerLeuHisAsnAlaGlnPheTyrLeuLeuGln  280
Db  781  GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCGCAATTTTATTTGCTACAA  840

Qy  281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuLeuIleMetAlaAla  300

Db  721  GAGATATTTCTCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGTGGGAAGGATCACC   780

Qy  261  AspSerHisGlnTrpAsnThrLeuAsnSerLeuHisAsnAlaGlnPheTyrLeuLeuGln  280
Db  781  GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCGCAATTTTATTTGCTACAA  840

Qy  281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuLeuIleMetAlaAla  300
Db  841  CGCACGCCAGAGGTGCCCGCACGCCGGCCACCCGGTTGATCATGGCGGCG           900

Qy  301  LeuThrProHisProProGlnLysGlnGlnAlaTyrGlyValThrLeuProThrSerValLeu  320
Db  901  TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTACTG   960

Qy  321  PheIleAlaGlyHisAspAThrAsnLeuAlaAsnLeuGlyGlyValAlaLeuGlyLeuAsnTrp  340
Db  961  TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGACTGGAGCTCAACTGG  1020

Qy  341  ThrLeuProGlyGlnProAspAsnThrProProGlyLeuGluValPheGluArgTrp  360
Db 1021  ACGCTTCCCGGTCAGCCCGATAACACGCCGCCAGGTGGTGAACTGGTGTTTGAACGGCTGG  1080

Qy  361  ArgArgLeuSerAspAsnSerGlnIleTrpIleGlnValSerPheGlnValPheGlnThrLeuGln  380
Db 1081  CGTCGGCTAAGCGATAACAGCCAGATGGATTCAGGTTTCCTGGTCTTCCAGACTTTACAG  1140

Qy  381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyLeuGluValProValLysLeuThr  400
Db 1141  CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCCCGGAGAGGTGAAACTGACC  1200

Qy  401  LeuAlaGlyLysCysGluGluValAspArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln  420
Db 1201  CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA  1260

Qy  421  IleValAsnGluGluAlaAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis  440
Db 1261  ATCGTGAATGAAGAAGCACGCGATACCGGCGTGCAGTTTGAGATCTCATCACCACCATCAC  1320
```

RESULT 5
US-09-866-379-1

; Sequence 1, Application US/09866379
; Patent No. 6855365
; GENERAL INFORMATION:
;   APPLICANT: DIVERSA CORPORATION
;   APPLICANT: SHORT, Jay
;   APPLICANT: KRETZ, Keith
;   APPLICANT: GRAY, Kevin
;   APPLICANT: BARTON, Nelson
;   APPLICANT: GARRETT, James
;   APPLICANT: O'DONOGHUE, Eileen
;   TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
;   FILE REFERENCE: DIVER1370-7
;   CURRENT APPLICATION NUMBER: US/09/866,379
;   CURRENT FILING DATE: 2001-05-24
;   PRIOR APPLICATION NUMBER: US 09/580,515
;   PRIOR APPLICATION FILING DATE: 2000-05-25
;   PRIOR APPLICATION NUMBER: US 09/318,528
;   PRIOR APPLICATION FILING DATE: 1999-05-25
;   PRIOR APPLICATION NUMBER: US 09/291,931
;   PRIOR APPLICATION FILING DATE: 1999-04-13
;   PRIOR APPLICATION NUMBER: US 09/259,214
;   PRIOR APPLICATION FILING DATE: 1999-03-01
;   PRIOR APPLICATION NUMBER: US 08/910,798
;   PRIOR APPLICATION FILING DATE: 1997-08-13
;   NUMBER OF SEQ ID NOS: 10
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1323
;   TYPE: DNA
;   ORGANISM: Escherichia coli
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1323)

```
                                                Best Local Similarity:   99.5%        Mismatches:    2
                                                Query Match:             97.1%        Indels:        0
                                                DB:                      3            Gaps:          0

                                     US-10-601-319-2 (1-440) x US-09-866-379-7 (1-1901)

Db   841 CGCACGCCAGAGGTTGCCGCCAGCCGCCCACCCCGTTATTGGATTTGGATCATGGCAGCG 900          Qy     1 MetLysAlaIleLeuIleLeuProPheLeuLeuLeuLeuProLeuProGlnSerAla  20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||               Db   188 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 247
Qy   301 LeuThrProHisAspPheProGlnLysGlnAlaThrTyrGlyValThrLeuProThrSerValLeu 320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||                    Qy    21 PheAlaGlnSerGluProGluLeuLysLeuSerValValIleValSerArgHisGly  40
Db   901 TTGACGGCCCATCCACCGCAAAAACAGGCGTATGGTGACATTACCCACTTCAGTACTG 960                Db   248 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGGTGATTGTCAGTCGTCATGGT 307

Qy   321 PheIleAlaGlyHisAspPheAsnLeuAlaLeuAsnLeuGlyTyrGlyValAlaLeuGluLeuAsnTrp 340     Qy    41 ValArgAlaProThrLysGlnAlaThrLeuMetGlnAspValThrProAlaTrpPro  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||                 Db   308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 367
Db   961 TTTATTGCCGGACACGATACTAATCTCGGCGGCGGAAATCTCGGCAATTCTGGAGCTCAACTGG 1020

Qy   341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyLeuValPheGluValArgTrp 360                Qy    61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu  80
         ||||||||||||||||||||||||||||||||||||||||||||||||||                           Db   368 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGGCGTGGTGAGCTAATCGCCTATCTC 427
Db  1021 ACGCTTCCCGGTCAGCCGGATAACACCCCGCCGGGTCGTGTTTGAACCGCTGG 1080

Qy   361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380                Qy    81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||                       Db   428 GGACATTACCAACGCCAGCGTCGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG 487
Db  1081 CGTCGGCTAAGCGATAACAGCCAGTGATTCAGGTTTCGCTGGTCTTCCAGACTTTTACAG 1140

Qy   381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400               Qy   101 GlnSerGlyGlnValAlaIleIleIleAlaAspValAspAspCysAlaIleThrValHis 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||                       Db   488 CAGTCTGGTCAGGTCGCGATTATTGCTGACGTTCGACGAGCGTACCCGTAAAACAGGCGAA 547
Db  1141 CAGATGCGTGATAAAAACGCCGCTGTCATTAAATACGCCCGCCGGAGGGTGAAAACTGACC 1200

Qy   401 LeuAlaGlyCysGluGluArgAsnAlaGlnMetCysSerLeuAlaGlyPheThrGln 420                  Qy   121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisGlnAlaAspThr 140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||                         Db   548 GCCTTCGCCGCCGGCGTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 607
Db  1201 CTGGCAGATGTGAAGAGCGAAATGCCAGGCCATGTTCGTTCGGCAGGTTTTTACCCAA 1260

Qy   421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHisHis 440               Qy   141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||                  Db   608 TCCAGTCCGATCCGTTATTTAATCCTCTAAAACTGGCTTTGCCAACTGGATAACGCG 667
Db  1261 ATCGTGAATGAAGCACGACAGCATACCGGGCGTGCCAGTTTGAGATCTCATCACCATCACCATCAC 1320

                                                                                      Qy   161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
                                                                                      Db   668 AACGTGACTGACGCGATCCTCAGCAGGCCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 727

                                                                                      Qy   181 ArgGlnThrAlaPheArgGluLeuGluValLeuAsnPheProGlnSerAsnLeuCys 200
                                                                                      Db   728 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 787

                                                                                      Qy   201 LeuLysArgGluValLysSerGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
                                                                                      Db   788 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC 847

                                                                                      Qy   221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
                                                                                      Db   848 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 907

                                                                                      Qy   241 GluIlePheLeuLeuGlnHisAlaGlnLeuMetProGlyLysTrpGlyArgIleThr 260
                                                                                      Db   908 GAGATATTTCTTCCTGCACAAGCACAGCAACTGATGCCGGGTAAGTGGGGAAGGATCACC 967

                                                                                      Qy   261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuAsnAlaGlnAsnPheTyrLeuLeuGln 280
                                                                                      Db   968 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA 1027

                                                                                      Qy   281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
                                                                                      Db  1028 CGCACGCCAGAGGTTGCCCGCAGCCGCCGCCACCCCGTTATTGGATTTGATCAAGACAGCG 1087

                                                                                      Qy   301 LeuThrProHisProProGlnLysGlnAlaThrTyrGlyValThrLeuProThrSerValLeu 320
                                                                                      Db  1088 TTGACGGCCCATCCACCGCAAAAACAGGCGTATGGTGACATTACCCACTTCAGTGCTG 1147

                                                                                      Qy   321 PheIleAlaGlyHisAspPheAsnLeuAlaAsnLeuGlyTyrGlyValAlaLeuGluLeuAsnTrp 340
                                                                                      Db  1148 TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGAGCTCAACTGG 1207
```

```
RESULT 6
US-09-866-379-7
; Sequence 7, Application US/09866379
; Patent No. 6855365
; GENERAL INFORMATION:
;  APPLICANT: DIVERSA CORPORATION
;  APPLICANT: SHORT, Jay
;  APPLICANT: KRETZ, Keith
;  APPLICANT: GRAY, Kevin
;  APPLICANT: BARTON, Nelson
;  APPLICANT: GARRETT, James
;  APPLICANT: O'DONOGHUE, Eileen
;  TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
;  FILE REFERENCE: DIVER1370-7
;  CURRENT APPLICATION NUMBER: US/09/866,379
;  CURRENT FILING DATE:   2001-05-24
;  PRIOR APPLICATION NUMBER: US 09/580,515
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: US 09/318,528
;  PRIOR FILING DATE: 1999-05-25
;  PRIOR APPLICATION NUMBER: US 09/291,931
;  PRIOR FILING DATE: 1999-04-13
;  PRIOR APPLICATION NUMBER: US 09/259,214
;  PRIOR FILING DATE: 1999-03-01
;  PRIOR APPLICATION NUMBER: US 08/910,798
;  PRIOR FILING DATE: 1997-08-13
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn version 3.1
;  SEQ ID NO 7
;  LENGTH: 1901
;  TYPE: DNA
;  ORGANISM: Escherichia coli
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (1)..(1901)
;  OTHER INFORMATION: n is any nucleotide
US-09-866-379-7

Alignment Scores:
Pred. No.:         2.34e-226      Length:          1901
Score:             2235.00        Matches:         430
Percent Similarity: 99.5%         Conservative:    0
```

Qy   21 PheAlaGlnSerGluProGluLeuLysSerValValIleValSerArgHisGly   40
Db 248 TTCGCTCAGAGTGAGCGCGAGCTGAAGTCGGTGGTGATAGTGTCAGTCGTCATGGT 307

Qy   41 ValArgAlaProThrLysAlaThrGlnMetGlnAspValThrProAspAlaTrpPro   60
Db 308 GTGCGTGCTCCAACCAAGGCCACGCAACTGCAGGATGTCACCCCAGACGCATGGCCA 367

Qy   61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGlyLeuIleAlaTyrLeu   80
Db 368 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC 427

Qy   81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuValAlaLysGlyLysCysPro 100
Db 428 GGACATTACCAACGCCAGCGTCGTGGCAGACGGATTGCTGGCGGAAAAAGGGCTGCCCG 487

Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspArgThrArgLysThrArgLysGlyGlu 120
Db 488 CAGTCTGGTTCAGGTCGCGGATTATTGCGGATGTCGACGAGCGTACCCGTAAAACAGGCGAA 547

Qy 121 AlaPheAlaAlaGlyLysLeuAlaProAspAspCysSerAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCTTCGCCGCCGGGCTGCCACCTGACTGTCAATAACCGTACATACCCAGGCAGATACG 607

Qy 141 SerSerProAspPheLeuPheAsnProLeuLysPheThrGlyValCysGlnLeuAspAsnAla 160
Db 608 TCCAGTCCGGATCCGTTATTTAATCCTCTAAAACTGGCGTTGCCAACTGGATAACGCG 667

Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyAlaGlySerIleGlySerPheThrGlyHis 180
Db 668 AACGTGACTGACGCGGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGCAT 727

Qy 181 ArgGlnThrAlaPheArgGluGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 728 CGGCAAACGGCGGTTCGGCAACTGGAACGGTGCTTAATTTTCCGCAATCAAACTTGTGC 787

Qy 201 LeuLysArgGluLysGlnAspSerGlnSerSerLeuThrGlnLeuAlaLeuProSerGluLeu 220
Db 788 CTTAAACGTGAGAAACAGGACAGAAGCTGTTCATTAACGCAGGCCATTACCATCGGAACTC 847

Qy 221 LysValSerAlaAspAspAsnValSerSerLeuThrGlyValAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTGAGCGCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCCATCAATGCTGACG 907

Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnMetProGluProGlyTrpGlyArgIleThr 260
Db 908 GAGATATTTTCTCCTGCAACAAGCACAGGAAATGCCGAGCCGGGGTGGGGAAGGATCACC 967

Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 968 GATTCACCACCAGTGGAACACCTTGCTAAGTTTGCATAACGCCCAATTTTGTTTGCTACAA 1027

Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 1028 CGGACGCCAGAGGTTGCCCAGCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 1087

Qy 301 LeuThrProHisProProGlnIleLysGlnAlaThrArgGlyAlaTyrArgGlyGluLeuValThrLeuProThrSerValLeu 320
Db 1088 TTGACGGCCCATCCACCGCAAAAACAGGCTATAGGTGTGACATTACCCACTTCAGTGCTG 1147

Qy 321 PheIleAlaGlyHisAspThrAsnLeuAsnLeuGlyAlaLeuGluLeuAsnTrp 340
Db 1148 TTTATCGCCGGACACGATACTAATCTCGGCAAATCTCGGCGCCACTGGAGCTCAACTGG 1207

Qy 341 ThrLeuProGlyGlnProAspAsnThrProGlyGlyGluLeuValPheGluArgTrp 360
Db 1208 ACGCTTCCCGGTCAGCCCGATAACACGCCCGGTGGTGAACTGGTGTTTGAACGCTGG 1267

Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1327

Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProGlyGlyGluValLysLysLeuThr 400
Db 1328 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCCGGAGAGGTGAAACTGACC 1387

Qy 401 LeuAlaGlyThrGlyGluGluArgGluAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1388 CTGGCAGGATGTGAAGAGCGAGAAAATGCGCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA 1447

Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTGAATGAATGAAGCACGCATACCGGCGTGCAGTTTG 1483

RESULT 7
US-09-866-379-9
; Sequence 9, Application US/09866379
; Patent No. 6855365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
;
; SEQ ID NO 9
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-9

Alignment Scores:
Pred. No.:   Length: 1901
Score: 2.1e-225   Matches: 429
                 2226.00
Percent Similarity: 99.3%   Conservative: 0
Best Local Similarity: 99.3%   Mismatches: 3
Query Match: 96.7%   Indels: 0
DB: 3   Gaps: 0

US-10-601-319-2 (1-440) x US-09-866-379-9 (1-1901)

Qy   1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla   20
Db 188 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 247

Qy   161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   662 AACGTGACTGACGCGATCCTCAGCGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 721

Qy   181 ArgGlnThrAlaPheArgGluLeuGlnSerLeuAsnPheProGlnSerAsnLeuCys 200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   722 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTTCCCCAATTAAACTTGTGC 781

Qy   201 LeuLysArgGluLysGlyGlnAspGluSerCysSerLeuThrGlnAlaAlaLeuProSerGluLeu 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   782 CTTAACCGTGAGAAACAGGAACGAAAGTGTTCATTAACCAGGCATTACCATCCGGAACTC 841

Qy   221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   842 AAGGTGAGCGGCCGACAATGTTTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 901

Qy   241 GluIlePheLeuLeuGlnInGlnAlaGlnGlyMetGluProGlyTrpProGlyTrpArgIleThr 260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   902 GAAATATTTTCTCCTGCACACAGGAATGCCGGAGCCCGGGGTGGGGAAGGATCACT 961

Qy   261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnAlaGlnPheTyrLeuLeuGln 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   962 GATTCACCACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTACTACAA 1021

Qy   281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1022 CGCACGCCAGAGGTTGCCCGCAGTCGGCCCACCCCGTTATTGGATTTGATCATGGCACG 1081

Qy   301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1082 TTGACGCCCCATCCACCGCAAAACAGGCGTATGGTGCACATTACCCACTTCAGTGCTG 1141

Qy   321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGlyLeuAsnLeuAsnTrp 340
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1142 TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGCCACTGGAGCTCAACTGG 1201

Qy   341 ThrLeuProGlyGlnProAspAsnProGlnGluLeuValPheGluLeuArgTrp 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1202 ACGCTTCCAGGTCCAGGTCAGGATAACACGCGCCAGGTGGTGAACGCTGG 1261

Qy   361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1262 CGTCGGCTAAGCGGATAACAGCCAGTGCAGTTCAGGTTTCGTCTGTCTTCCAGACTTTACAG 1321

Qy   381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1322 CAGATGCGTGATAAAACGCCGCTATCATTAAATACGCCGCCCGGAGAGGTGAAACTGACC 1381

Qy   401 LeuAlaGlyCysGluGluArgArgAsnAlaGlnInGlyMetCysSerSerLeuAlaGlyPheThrGln 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1382 CTGGCAGGATGTGAAGAGCGAAATGCCCAGGGCCATGTGTTCGTTGGCCGGTTTTACGCAA 1441

Qy   421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
           ||||||||||||||||||||||||||||||||||||
Db   1442 ATCGTGAATGAAGCGGCGCATACCGGCGCAGTTTG 1477

Db   1328 CAGATGCGTGATAAAACGCCGTGTCATTAAATAACCGCCCGGAGAGGTGAAACTGACC 1387
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   401 LeuAlaGlyCysGluGluArgArgAsnAlaGlnInGlyMetCysSerSerLeuAlaGlyPheThrGln 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388 CTGGCAGGATGTGAAGAGCGAAATGCCAGGGCCATGTGTTCGTTGGCCGGTTTTACGCAA 1447

Qy   421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
           ||||||||||||||||||||||||||||||||||||
Db   1448 ATCGTGAATGAAGCACGCGCATACCGGCGCAGTTTG 1483

Alignment Scores:
Pred. No.:   Length: 1489
Score: 2221.00   Matches: 428
Percent Similarity: 99.1%   Conservative: 0
Best Local Similarity: 99.1%   Mismatches: 4
Query Match: 96.5%   Indels: 0
DB: Gaps: 0
                                                     3

US-10-601-319-2 (1-440) x US-09-540-149A-9 (1-1489)

Qy   1   MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   182 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTTTGATTCCGTTAACCCCGCAAATCTGCA 241

Qy   21   PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   242 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGTCGGTAGTTGTCAGCCGTCATGGT 301

Qy   41   ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpProPro 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   302 GTGCGTGCCCCACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGGCCATGCCA 361

Qy   61   ThrTrpProValLysTrpGlyGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   362 ACCTGGCCGGTAAAACTGGGCTGGTGGCTGACACCACGCGGGGTGTGAGCTAATCGCCTATCTC 421

Qy   81   GlyHisTyrGlnInArgGlnAsnArgLeuValAlaAspProGlyLeuLeuAlaLysGlyCysProPro 100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   422 GGACATTACCAACGCCAGCCGTCTGGTGGCCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG 481

Qy   101 GlnSerGlyInValIleAlaIleAlaAspAlaAspValAspArgThrArglysThrGlyGlu 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   482 CAGCCGTGGTCAGGTCGCGATTATTGCTGCGACGCGACGGTACCCGTAAAACAGGCGAA 541

Qy   121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   542 GCCTTCGCCGCCGGCTGGCACCTGACTGTCAATAACCGTACATACCCCAGGCAGATACG 601

Qy   141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   602 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG 661

```
;  LENGTH: 1489
;  TYPE: DNA
;  ORGANISM: Escherichia coli
US-10-266-041A-9

Alignment Scores:
Pred. No.:              4.81e-225
Score:                  2221.00
Percent Similarity:     99.1%
Best Local Similarity:  99.1%
Query Match:            96.5%
DB:                     4

US-10-601-319-2 (1-440) x US-10-266-041A-9 (1-1489)

Qy    1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  182 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTTTGATTCCGTTAACCCCGCAATCTGCA 241

Qy   21 PheAlaGlnSerGluProGluLeuLysGluValValIleValSerArgHisGly 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  242 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGTGGTGATTGTCAGCCGTCATGGT 301

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  302 GTGCGTGCCCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 361

Qy   61 ThrTrpProValLysGlyGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  362 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCACGCGGTGGCGAGCTAATCGCCTATCTC 421

Qy   81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  422 GGACATTACCAACGCCAGCGTCTCGGTGGCCGACCGATTCTCGGCGAAAAAGGGCTGCCCG 481

Qy  101 GlnSerGlyGlnValAlaIleLeuAlaAspValAspGluArgThrArgLysThrGlyGlu 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  482 CAGCCTGGTCAGGTCGCGATTATTGCTGACGTCGACGAGCGTACCCGTAAAACAGGCGAA 541

Qy  121 AlaPheAlaAlaGlyLysLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  542 GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 601

Qy  141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAspAsnAla 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  602 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACCGGCGTTTGCCAACTGGATAACGCG 661

Qy  161 AsnValThrAspAlaAlaIleLeuSerArgAlaGlyGlySerIleAlaAlaAspPheThrGlyHis 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  662 AACGTGACTGACGCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 721

Qy  181 ArgGlnThrArgAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
          ||||||||||||||||||||||||||||||||  ||| ||||||||||||||||||||||
Db  722 CGGCAAACGCGGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTTCCCAATTAAACTGTGC 781

Qy  201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnLeuAlaLeuProSerGluLeu 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  782 CTTAACCGTGAGAAACAGGACGAAAGCTGTTCATTAACCCAGCTGGCATTACCATCCGAACTC 841

Qy  221 LysValSerAlaAspAsnAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  842 AAGGTGAGCGCCGACAATGTTTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 901

Qy  241 GluIlePheLeuLeuGlnGlnAlaGlnGlnMetProGluAlaGlnGlyMetCysProGlyTrpProGlyTyrArgIleThr 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  902 GAAATATTTTCTCCTGCAACAAGCACAGCAGATGCCGGAAGCGCAGGGGATGTGGGGAAGGAAGGATCACT 961

Qy  261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  962 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTACTACAA 1021

Qy  281 ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 1022 CGCACGCCAGAGGTTGCCGCACGGTCGCGCACCCCGTTATTGGATTTGATCATGGCAGCG 1081

Qy  301 LeuThrProHisProProGlnLysGlnAlaAlaTyrGlyValValThrGlyValLeuProThrSerValLeu 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1082 TTGACGCCCCATCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCTG 1141

Qy  321 PheIleAlaGlyHisAspAlaGlyAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 340
          |||||||||  |||||||||||||||||||||||||||||||||||||||||||||||||
Db 1142 TTTATTGCCGACACGATACTCATCTGGCAAATCTCCGGCGCCACTGGAGCTCAACTGG 1201

Qy  341 ThrLeuProGlyGlnInProAspAsnThrProProGlyGlyLeuValPheGluArgTrp 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1202 ACGCTTCCAGTCAGCCGGATAACACGCCCCCAGGTGGTGAACTGGTGTTTGAACCCTGG 1261

Qy  361 ArgArgLeuSerArgAspAsnSerGlnInTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1262 CGTCGGCTAAGCAGCGGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1321

Qy  381 GlnMetArgAspLysTyrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1322 CAGATGCGTGATAAAAACGCCGCTATCATTAATACGCCGCCCGGAGAGGTGAAACTGACC 1381

Qy  401 LeuAlaGlyCysGluGluArgAsnAlaGlnInGlyMetCysSerLeuAlaGlyPheThrGln 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1382 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCGTTGGCCGGGTTTTACGCAA 1441

Qy  421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
          ||||||||||||||||||||||||||||||||||||
Db 1442 ATCGTGAATGAAGCGGCGATACCGGTGCAGTTTG 1477
```

```
RESULT 10
US-09-866-379-6
; Sequence 6, Application US/09866379
; Patent No. 6855365
; GENERAL INFORMATION:
;  APPLICANT: DIVERSA CORPORATION
;  APPLICANT: SHORT, Jay
;  APPLICANT: KRETZ, Keith
;  APPLICANT: GRAY, Kevin
;  APPLICANT: BARTON, Nelson
;  APPLICANT: GARRETT, James
;  APPLICANT: O'DONOGHUE, Eileen
;  TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
;  FILE REFERENCE: DIVER1370-7
;  CURRENT APPLICATION NUMBER: US/09/866,379
;  CURRENT FILING DATE: 2001-05-24
;  PRIOR APPLICATION NUMBER: US 09/580,515
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: US 09/318,528
;  PRIOR FILING DATE: 1999-05-25
;  PRIOR APPLICATION NUMBER: US 09/291,931
;  PRIOR FILING DATE: 1999-04-13
;  PRIOR APPLICATION NUMBER: US 09/259,214
;  PRIOR FILING DATE: 1999-03-01
;  PRIOR APPLICATION NUMBER: US 08/910,798
;  PRIOR FILING DATE: 1997-08-13
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
;  LENGTH: 1901
;  TYPE: DNA
;  ORGANISM: Escherichia coli
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (1)..(1901)
;  OTHER INFORMATION: n is any nucleotide
US-09-866-379-6

Alignment Scores:
Pred. No.:                          Length:              1901
Score:             7.09e-225        Matches:             428
                   2221.00          Conservative:        0
Percent Similarity: 99.1%           Mismatches:          4
Best Local Similarity: 99.1%
```

Query Match: 96.5%   Indels: 3   0
DB: Gaps: 3   0

US-10-601-319-2 (1-440) x US-09-866-379-6 (1-1901)

```
Qy    1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  188 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 247

Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  248 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGGTGGTCAGTCGTCATGGT 307

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 367

Qy   61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
        ||||| | |||||||||||||||||||||||||||||||||||||||||||||
Db  368 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCCGCNGGTGGTGAGCTAATCGCCTATCTC 427

Qy   81 GlyHisTyrGlnArgGlnArgGlyLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
        |||||||||||    |||||||||||||||||||||||||||||||||||
Db  428 GGACATTACCTGGGCCCAGGCGTCTGGTAGCCGGATTGCTGGCGAAAAAGGGCTGCCCG 487

Qy  101 GlnSerGlyIleValAlaIleIleLeuAlaAspGluArgThrArgLysThrGlyGluGlu 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  488 CAGTCTGGTCAGTGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA 547

Qy  121 AlaPheAlaAlaAlaGlyLeuAlaProCysAlaIleThrValHisThrGlnAlaAspThr 140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  548 GCCTTCGCCGCCGGGCTGGCACCTGACGTGTGCAATAACCGTACATACCCAGGCAGATACG 607

Qy  141 SerSerProAspAspLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  608 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG 667

Qy  161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  668 AACGTGACTGACGCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 727

Qy  181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  728 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTGTGC 787

Qy  201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnIleAlaLeuProSerGluLeu 220
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  788 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGCATTACCATCGGAACTC 847

Qy  221 LysValSerAlaAspAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  848 AAGGTGAGCGCGGACGACAATGTCTCATTAACCGGTGCCGTAAGCCTCGCATCAATGCTGACG 907

Qy  241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluGlyTrpGlyArgIleThr 260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  908 GAGATATTTCTCCTGCAACAGCAACAGGGAATGCCGGAGCGGGGTGGGGAAGGATCACC 967

Qy  261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  968 GATTCACCAGTGGAACACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA 1027

Qy  281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1028 CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 1087

Qy  301 LeuThrProHisProProGlnLysProGlnGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1088 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCTG 1147

Qy  321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1148 TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCCACTGGAGCTCAACTGG 1207
```

(right column)

```
Qy  341 ThrLeuProGlyGlnProAspAsnThrProGlyGlyGluLeuValPheGluArgTrp 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1208 ACCTTCCCGGTCAGCCGGTCAGATAACACGCCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1267

Qy  361 ArgArgLeuSerAspAsnSerGlnTrpIleValSerLeuAsnValPheGlnThrLeuGln 380
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1268 CGTCGGCTAAGCGATAACAGCCAGTGGATTGTTTCGCTGGTCTTCCAGACTTTACAG 1327

Qy  381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1328 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCCGGAGAGGTGAAACTGACC 1387

Qy  401 LeuAlaGlyCysGluGluArgAsnAlaGlnLeuMetCysSerLeuAlaGlyPheThrGln 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1388 CTGGCAGGATGTGAAGAGCGAAATGCCAGGCCATGTGTTCGTTGGCAGGTTTTACGCAA 1447

Qy  421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
        ||||||||||||||||||||||||||||||||||||
Db 1448 ATCGTGATGAATGAAGCACGCATACCGGCGTGCAGTTTG 1483
```

RESULT 11
US-09-866-379-5
; Sequence 5, Application US/09866379
; Patent No. 6855365
; GENERAL INFORMATION:
;   APPLICANT: DIVERSA CORPORATION
;   APPLICANT: SHORT, Jay
;   APPLICANT: KRETZ, Keith
;   APPLICANT: GRAY, Kevin
;   APPLICANT: BARTON, Nelson
;   APPLICANT: GARRETT, James
;   APPLICANT: O'DONOGHUE, Eileen
;   TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
;   FILE REFERENCE: DIVER1370-7
;   CURRENT APPLICATION NUMBER: US/09/866,379
;   CURRENT FILING DATE: 2001-05-24
;   PRIOR APPLICATION NUMBER: US 09/580,515
;   PRIOR FILING DATE: 2000-05-25
;   PRIOR APPLICATION NUMBER: US 09/318,528
;   PRIOR FILING DATE: 1999-05-25
;   PRIOR APPLICATION NUMBER: US 09/291,931
;   PRIOR FILING DATE: 1999-04-13
;   PRIOR APPLICATION NUMBER: US 09/259,214
;   PRIOR FILING DATE: 1999-03-01
;   PRIOR APPLICATION NUMBER: US 08/910,798
;   PRIOR FILING DATE: 1997-08-13
;   NUMBER OF SEQ ID NOS: 10
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
;   LENGTH: 1901
;   TYPE: DNA
;   ORGANISM: Escherichia coli
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1901)
;   OTHER INFORMATION: n is any nucleotide
US-09-866-379-5

Alignment Scores:
Pred. No.:   Length: 1901
Score: 9.05e-225   Matches: 429
                    2220.00   Conservative: 0
Percent Similarity: 99.3%   Mismatches: 3
Best Local Similarity: 99.3%   Indels: 0
Query Match: 96.4%   Gaps: 0
DB: 3

US-10-601-319-2 (1-440) x US-09-866-379-5 (1-1901)

```
Qy    1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  188 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 247

Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
```

```
Db   248  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TTCGCTCAGAGTGAGCCCGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT  307

Qy    41  ValArgAlaProThrLysAlaThrLeuMetGlnHisAspValThrProAlaTrpPro      60
Db   308  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GTGCGTGCTCCAACCAAGGCCACGCCAACTGATGCAGGAGGATGTCACCCCAGACGCATGGCCA  367

Qy    61  ThrTrpProValLysLeuGlyTrpProLeuThrProArgGlyGlyLeuIleAlaTyrLeu    80
Db   368  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          ACCTGGCCGGTAAAACTGGGTTGACACCGCGNGTGGTGACCAGTAATCGCCTATCTC  427

Qy    81  GlyHisTyrGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro         100
Db   428  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GGACATTACCAAGCGCGACGGATTCGTCGGTGGCGACAGACGCTGCGGCCCG  487

Qy   101  GlnSerGlyValAlaIleIleAlaAspValAspThrArgLysThrGlyGlu           120
Db   488  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CAGTCTGGTCAGGTCGGCGATTATTGCTGAGTCGACGAGCGTACCCGTAAAACAGGGCGAA  547

Qy   121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr   140
Db   548  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GCCTTCGCCGCCGGGCTGGCACCTGACATGTCAATAACCGTACATACCCAGGCAGATACG  607

Qy   141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla   160
Db   608  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG  667

Qy   161  AsnValThrAspAlaAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis  180
Db   668  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          AACGTGACTGACGCGATCCTCAGCAGGGCCAGGAGGGTCAATTGCTGACTTTTACCGGGCAT  727

Qy   181  ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys    200
Db   728  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CGGCAAACGGCGTTTCGGCAACTGGAACGGGTCGTCTTAATTTTCCGCAATCAAACTTGTGC  787

Qy   201  LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu    220
Db   788  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CTTAAACGTGAGAAACAGGATCGGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC  847

Qy   221  LysValSerAlaAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr  240
Db   848  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG  907

Qy   241  GluIlePheLeuLeuGlnGlnAlaGlnIleGlyMetProGlyLeuTyrTrpGlyArgIleThr  260
Db   908  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GAGATATTTTCTCTGCACAAGCACAGGGAATCGGAGCCCGGAGCGGGTGGGGAAGGATCACC  967

Qy   261  AspSerHisGlnTrpAsnThrLeuTrpSerLeuAsnAlaGlnPheTyrLeuLeuGln        280
Db   968  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTTATTTGCTACAA  1027

Qy   281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLeuMetAlaAla  300
Db  1028  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CGCACGCCAGAGGTTGCCCGCAGCCGCCGCCACCCCGTTATTAGATTTGATCAAGACACGCG  1087

Qy   301  LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu    320
Db  1088  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TTGACGCCCCATCCACCCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCTGT  1147

Qy   321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp    340
Db  1148  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TTTATCGCCGGACACGATTAATACTGCAAATCTCGGCGCGGCCACTCGGAGCTCAACTGG  1207

Qy   341  ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlyLeuValPheGluArgTrp     360
Db  1208  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          ACGCTTCCCGGTCAGCCGGATAACACGCCCGGTGGTGGGTTGGTGTTTGAACGCTGG  1267

Qy   361  ArgArgLeuSerAspAsnSerGlnTrpIleGlnTrpLeuSerLeuValPheGlnThrLeuGln  380
Db  1268  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGTGGCTGTCGCTGGTCTTTCAGACTTTACAG  1327

Qy   381  GlnMetArgAspLysAspThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr  400
Db        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  1328  CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCCGGAGAGGTGAAACTGACC  1387

Qy   401  LeuAlaGlyCysGluGluArgAsnAlaGlnAsnAlaGlyMetCysSerLeuAlaGlyPheThrGln  420
Db  1388  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CTGGCAGGATGTGAAGAGCGAAATGCGCAGGAGCATGGTGTTCGTTGGCAGTTTTACGCAA  1447

Qy   421  IleValAsnGluAlaLysArgIleProAlaCysSerLeu  432
Db  1448  |||||||||||||||||||||||||||||||||||||||
          ATCGTGAATGAAGCACGCATACCCGGCGTGCAGTTTG  1483
```

RESULT 12
US-09-715-477-2
; Sequence 2, Application US/09715477
; Patent No. 6841370
; GENERAL INFORMATION:
;   APPLICANT: Lei, Xingen
;   TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
;   FILE REFERENCE: 19603/4031
;   CURRENT APPLICATION NUMBER: US/09/715,477
;   CURRENT FILING DATE: 2000-11-17
;   PRIOR APPLICATION NUMBER: 60/166,179
;   PRIOR FILING DATE: 1999-11-18
;   NUMBER OF SEQ ID NOS: 9
;   SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 2
;   LENGTH: 1486
;   TYPE: DNA
;   ORGANISM: Escherichia coli
US-09-715-477-2

Alignment Scores:
Pred. No.:   Length: 1486
Score: 2217.00   Matches: 427
Percent Similarity: 98.8%   Conservative: 0
Best Local Similarity: 98.8%   Mismatches: 5
Query Match: 96.3%   Indels: 0
DB: Gaps: 0

US-10-601-319-2 (1-440) x US-09-715-477-2 (1-1486)

```
Qy     1  MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla    20
Db   188  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          ATGAAAGCGATCTTAATCCCATTTTTATCTCTGATTCCGTTAACCCCGCAATCTGCA  247

Qy    21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly    40
Db   248  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TTCGCTCAGAGTGAGCCCGAGCTGAAAGTGTGGAAAGTGTGATGATTGTCAGTCGTCATGGT  307

Qy    41  ValArgAlaProThrLysAlaThrLeuMetGlnHisAspValThrProAlaTrpProAla    60
Db   308  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GTGCGTGCTCCAACCAAGGCCACGCCAACTGATGCAGGAGGATGTCACCCCAGACGCATGGCCA  367

Qy    61  ValArgAlaProThrLysAlaThrLeuMetGlnHisAspValThrProAlaTrpProAspAlaTrpLeu  80
Db   368  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GTGCGTGCCCCAACCAAGGCCACCAAGGCTCGGATCGTGGACGCAGGTCACCCCCAGACGCG  427

Qy    81  GlyHisTyrGlnArgLeuValAlaArgLeuValAlaAlaAspArgGlyLeuLeuAlaLysLysGlyCysPro  100
Db   428  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GGACATTACCAAGCGCGCCAGCGGTTGCCGCGAAGAAGAAGGGGCTGCCGAAAAAGGGCTGCCCG  487

Qy   101  GlnSerGlyValAlaIleIleAlaAspValAspThrArgLysThrGlyGlu           120
Db   488  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CAGTCTGGTCAGGTCGGCGATTATTGTGAAGTGTGGAAAGTGTGATTGTCAGCCGTCATGGT  307

Qy   121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspTrpProAla  140
Db   548  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATTGTGTCACCCGTAAAACAGGGCCAA  607

Qy   141  ValArgAlaProThrLysAlaThrLeuMetGlnHisAspValThrProAlaTrpProAspAla  160
Db   608  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GTGCGTGCCCCAACCAAGGCCACCAAGGCTCGGATCGTGGACGCAGGTCACCCCAGACGCG  367

Qy   161  ThrTrpProValLysLeuGlyTrpProLeuThrProArgGlyGlyLeuAlaTyrLeu     80
Db   368  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          ACCTGGCCGGTAAAACTGGGTTGGCCGCTGACCACCAGCGGGGTGGGTGAGCTAATCGCCTATCTC  427

Qy    81  GlyHisTyrGlnArgLeuValAlaArgLeuValAlaAlaAspArgGlyLeuLeuAlaLysLysGlyCysPro  100
Db   428  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GGACATTACCAAGCGCGCCAGCGGTTGCCGCGAAGAAGAAGGGGCTGCCGAAAAAGGGGCTGCCCG  487

Qy   101  GlnSerGlyValAlaIleIleAlaAspValAspThrArgLysThrGlyGlu           120
Db   488  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CAGTCTGGTCAGGTCGGCGATTATTGTGAAGTGTGGAAAGTGTGATTGTCAGCCGTCATGGT  307

Qy   121  AlaPheAlaAlaGlyLeuAlaProAspCysSerAlaIleThrValHisThrGlnAlaAspTrpProAla  140
Db   548  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GCCTTCGCCGCCGGGCTGGCACCCGGGTCGCCGGATAATTGTGCAATTGGTGTCACCCCGTAAAACAGGACGATACG  607

Qy   141  SerSerProAspProLeuPheAsnProLeuLysCysGlnLeuAspAsnAla           160
Db   608  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GCCTTCGCCGCCGGGCTGGCACCTGGCTGTGCATGTCAATAACCTGTACATACCCAGGCAGATATACG  607

Qy   161  AsnValThrAspAlaAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis  180
Db        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

US-09-715-477-4

Alignment Scores:
Pred. No.:
Score: 2197.00
Percent Similarity: 98.6%
Best Local Similarity: 98.1%
Query Match: 95.4%
DB:

Length: 1486
Matches: 424
Conservative: 2
Mismatches: 6
Indels: 0
Gaps: 0

US-10-601-319-2 (1-440) x US-09-715-477-4 (1-1486)

```
Qy   1  MetLysAlaIleLeuIlePheProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla  20
Db 188  ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA  247

Qy  21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly  40
Db 248  TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGCCGTCATGGT  307

Qy  41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro  60
Db 308  GTGCGTGCCCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA  367

Qy  61  ThrTrpProValLysGlyValGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu  80
Db 368  ACCTGGCCGGTAAAACTGGGTTGGCTGACACCACGCGGTGGTGAGCTAATCGCCTATCTC  427

Qy  81  GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 428  GGACATTACCAGCGCCAGCGTCTGGTGGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG  487

Qy 101  GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 488  CAGCCTGGTCAGGTCGCCGATTATTGTCGATGTCGACGAGCGTACCATACCAGGCAGATACG 547

Qy 121  AlaPheLeuAlaGlyLeuLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548  GCCTTCCTGGCCGGGCTGCTGGCACCTGACTGTGCAAATAACCAGGCAGATACG       607

Qy 141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 608  TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG  667

Qy 161  AsnValThrAspAlaAlaLeuSerArgAlaGlyLysSerIleAlaAspPheThrGlyHis 180
Db 668  AACGTGACTGACGCCGCCCTGAGCCGCGCCGGCAAATCGATTGCTGACTTTACCGGGCAT  727

Qy 181  ArgGlnThrAlaPheArgGluLeuValLeuAsnPheProGlnSerLeuCys 200
Db 728  CGGCAAACGGCGTTTCGCGAACTGGTGCTTAATTTTCCGCAATCAAACTTGTGC       787

Qy 201  LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 788  CTTAAACGTGAGAAACAGGACGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC  847
```

```
Db 668  AACGTGACTGACGGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT  727

Qy 181  ArgGlnThrAlaPheArgGluLeuValLeuAsnPheProGlnSerLeuCys 200
Db 728  CGGCAAACGGCGTTTCGCAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC  787

Qy 201  LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 788  CTTAAACGTGAGAAACAGGACGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC 847

Qy 221  LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 848  AAGGTGAGCGCCGACAATGTTTCATTAACCGGTGCGGTAAGCCTGGCCATCAATGCTGACG 907

Qy 241  GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGlyLeuTrpGlyArgIleThr 260
Db 908  GAAATATTTTCCTGCAACAGGCACAGGGAATGCCCGGACCGTGGGGAAGGATCACT  967

Qy 261  AspSerHisGlnTrpAsnThrLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 968  GATTCACACCAGTGGAACACCCTGTCAAGTTTGCATAACGCGCAATTTTATTACTACAA 1027

Qy 281  ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 1028 CGCACGCCAGAGGTTGCCCGCAGTCGCGCCACCCCGTTATTGGATTTGATCAAGAGACGCG 1087

Qy 301  LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 1088 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCACTTCAGTGCTG 1147

Qy 321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 1148 TTTATTGCCGGACACGATACTAATCTCGGCGCGCCAGCTGGGGAGCTGAAACTGGTGG 1207

Qy 341  ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1208 ACGCTTCCAGGTCAGCCCGGATAACACGCCGCCAGGTGGTGAACTGGTGTTTGAACCGTGG 1267

Qy 361  ArgArgLeuSerAspAspSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 CGTCGGCTAAGCGATGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1327

Qy 381  GlnMetArgAspLysProThrLeuSerAsnThrProProGlyGluValLysLeuProThr 400
Db 1328 CAGATGCGTGATAAAACCGCCCTATCATTAATACGCCGCCAGGTGGTGAAACTGACC 1387

Qy 401  LeuAlaGlyCysGluGluLysArgAsnAlaGlnGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1388 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCCTTCGTTGGCCGGTTTTACGCAA 1447

Qy 421  IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTGAATGAAGCGCGCATACCGGCGTGCAGTTTG 1483
```

RESULT 13

```
Pred. No.:                              Length:            1272
Score:                    2.39e-222     Matches:           423
Percent Similarity:       2194.50       Conservative:      0
Best Local Similarity:    96.1%         Mismatches:        0
Query Match:              96.1%         Indels:            17
DB:                       95.3%         Gaps:              1
                          2

US-10-601-319-2 (1-440)  x  US-08-910-798-1 (1-1272)

Qy      1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla    20
                                                                          60
Db      1 ATGAAAGCGATCTAATCCCATTTTATCTCTTCTGATTCCTTAACCCGCCAATCTGCA

Qy     21 PheAlaGlnSerGluProGluLeuLysLeuSerValIleValSerArgHisGly          40
          |||||||||||||||||||||||||||||||||||||||||||||||||||||         120
Db     61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGATGTCAGTCGTCATGGT

Qy     41 ValArgAlaProThrLysAlaThrGlnLeuMetGluAspValThrProAspAlaTrpPro     60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||     180
Db    121 GTGGGTGCTCCAACCAAGGCCACCCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA

Qy     61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu     80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||     240
Db    181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC

Qy     81 GlyHisTyrGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyLeuLysGlyCysPro    100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||    300
Db    241 GGACATTACCAACGCAGCGTCTCGTAGCCACGGATTGCTGGCGAAAAGGGCTGCCCG

Qy    101 GlnSerGlyGlnValAlaIleIleAlaIleAlaAspValAlaAspValAlaArgLysThrArgLysThrGlyGlu    120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||    357
Db    301 CAGTTCTGGTCAGGTGCCGGATTATTGCTGAGTGTCGACGAGCGTACCCGTAAAACAGGC---

Qy    121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr    140
          |||||||||||||||||||||||||||||||||                               369
Db    358 -------------------------------------CAGGCAGAGATACG

Qy    141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla    160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||    429
Db    370 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAACTGGGCGTTTGCCAACTGGATAACGCG

Qy    161 AsnValThrAspAlaAlaIleLeuSerArgAlaArgAlaGlyGlySerIleAlaAspPheThrGlyHis    180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||    489
Db    430 AACGTGACTGACCGCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT

Qy    181 ArgGlnThrAlaPheArgGluLeuGluAlaArgValLeuAsnPheProGlnSerAsnLeuCys    200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||    549
Db    490 CGGCAAACGGCGTTTCGGGAACTGGAACGGGTGCTTAATTTCCGCAATCAAACTGTGTGC

Qy    201 LeuLysArgGluLysGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu          220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||         609
Db    550 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCCAGGCATTACCATCCGGAACTC

Qy    221 LysValSerAlaAspAsnAlaValSerLeuThrGlyAlaAlaValAlaSerLeuSerMetLeuThr    240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||    669
Db    610 AAGGTGAGCGCCGACAATGTCCATTAACCGGTGCGCGCATCAATGCTGACG

Qy    241 GluIlePheLeuLeuGlyIleAsnGlnAlaGlnHisGluLeuMetProGluProGlyTyrArgIleThr    260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||    729
Db    670 GAGATATTTTCTCCTGCAACAGCACAGGAAATGCCAGGGACGCGGGGTGGGAAGGATCACC

Qy    261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuAsnAlaGlnInPheTyrLeuLeuGln    280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||    789
Db    730 GATTCACACCAGTGGAACACCTTGCTAAGTTGCATAACGCGCAATTTTATTTGCTACA

Qy    281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla    300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||    849
Db    790 CGCACGCCAGAGTTGCCCGCAGCCGCCACCCCGTTATTGGATTTGATCATGGCAGCG

Qy    301 LeuThrProHisProGlnLysGlnLysAlaThrTyrGlyValThrLeuProThrSerValLeu    320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||    909
Db    850 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTACTG
```

```
Db   1088 TTGACGCGCCCCATCCACCGCAAAAACAGGCGTATGGTGGTGACATTACCCACTTCAGTGCTG   1147
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp    340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148 TTTATTGCGGGACACGACACTAATCTGGCAAATCTCGGCGGCGCCACTGAGCTCAACTGG   1207

Qy    341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluValPheGluGlnArgTrp   360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208 ACGCTTCCAGGTCAGCCGGATAACACGCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG   1267

Qy    361 ArgArgLeuSerAspAsnSerGlnIleGlnIleTrpIleGlnValSerLeuValPheGlnThrLeuGln    380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268 CGTCGGCTAAGCGACGATAACAGCCCAGTGGATTCAGTTTCGTCCTGGTCTTCCAGACTTTACAG   1327

Qy    381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValValLysLeuThr    400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328 CAGATGCGTGATAAAACGCCGCTATCATTAAATACGCCGCCGGGAGAGGTGAAACTGACC   1387

Qy    401 LeuAlaGlyCysGluGluGlnArgAsnAlaGlnGlyMetGlyMetCysSerLeuAlaGlyPheThrGln    420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388 CTGGCAGGATGTGAAGAGCGAAATGCCAGGGCATGTGTTCGTTGGCCGGTTTACGCAA   1447

Qy    421 IleValAsnGluAlaArgIleProAlaCysSerLeu    432
          |||||||||||||||||||||||||||||||||||||
Db   1448 ATCGTGAATGAATGAAGCGCGCATACCGGCGTGCAGTTTG   1483
```

```
RESULT 14
US-08-910-798-1

; Sequence 1, Application US/08910798
; Patent No. 5876997
; GENERAL INFORMATION:
;  APPLICANT:  KRETZ
;  TITLE OF INVENTION:   NOVEL PHYTASE
;  NUMBER OF SEQUENCES:  4
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE:  Fish & Richardson P.C.
;   STREET:  4225 Executive Square, Suite 1400
;   CITY:  La Jolla
;   STATE:  California
;   COUNTRY:  US
;   ZIP:  92037
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE:  Floppy disk
;   COMPUTER:  IBM PC compatible
;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:  PatentIn Release #1.0, Version #1.25
;  CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/910,798
;   FILING DATE:  August 13,1997
;   CLASSIFICATION:  435
;  ATTORNEY/AGENT INFORMATION:
;   NAME:  HAILE, PH.D., LISA A.
;   REGISTRATION NUMBER:  38,347
;   REFERENCE/DOCKET NUMBER:  09010/029001
;  TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  619/678-5070
;   TELEFAX:  619/678-5099
;  INFORMATION FOR SEQ ID NO:  1:
;   SEQUENCE CHARACTERISTICS:
;    LENGTH:  1272 base pairs
;    TYPE:  nucleic acid
;    STRANDEDNESS:  double
;    TOPOLOGY:  linear
;   MOLECULE TYPE:  DNA (genomic)
;   IMMEDIATE SOURCE:
;    CLONE:  PHYTASE
;   FEATURE:
;    NAME/KEY:
;    LOCATION:
US-08-910-798-1
Alignment Scores:
```

```
Qy    321  PheIleAlaGlyHisAspThrLeuAlaAsnLeuGlyGlyAlaAlaLeuGluLeuAsnTrp  340
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    910  TTTATTGCCGGACACGATACTACTTGGCAAATCTGGCGGCGCCACTGGAGCTCAACTGG  969

Qy    341  ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp  360
           ||||||||||||||||||||||||||||||||||||||||||||||
Db    970  ACGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG  1029

Qy    361  ArgArgLeuSerAspAsnSerGlnTrpIleGlnPheGlnThrLeuGln  380
           ||||||||||||||||||||||||||||||||||||||||||||
Db   1030  CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGTTTCGCTGGTCTTCCAGACTTTACAG  1089

Qy    381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr  400
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1090  CAGATGCGTGATAAAACCCCGCTGTCATTAAATACGCCGCCCGGAGAGGTGAAACTGACC  1149

Qy    401  LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln  420
           |||||||||||||||||||  |||||  |||||||||||||||||||||
Db   1150  CTGGCAGGATGTGAAGAGCGCAGGGCATGGTTCGTTGGCAGGTTTTACGCAA  1209

Qy    421  IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHisHis  440
           |||||||||||||||||||||||||||||||
Db   1210  ATCGTGAATGAAGCACCCATACCGGCCGTGCAGTTTGAGATCTCATCACCATCACCATCAC  1269
```

RESULT 15
US-09-489-039A-341
; Sequence 341, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 341
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-341

Alignment Scores:
```
Pred. No.:          6.88e-46          Length:          1266
Score:              528.00            Matches:         134
Percent Similarity: 48.7%             Conservative:    69
Best Local Similarity: 32.1%          Mismatches:      184
Query Match:        22.9%             Indels:          30
                                      Gaps:            10
DB:                 3
```

US-10-601-319-2 (1-440) x US-09-489-039A-341 (1-1266)

```
Qy     19  SerAlaPheAlaGlnSerGlu--------ProGlu--LeuLysLeuSerVal  33
           |||||||||||:||::|||::      ||||| ::|||||:|||||||||
Db     67  TCCGCTGGCGCTCAGGCGCAGGACAAGGCCCCGAGGGGTATCAGCTGCAGCAGGTG  126

Qy     34  ValIleValSerArgHisGlyValArgAlaPro---ThrLysAlaThrGlnLeuMetGln  52
           ||||||||::||||||||||||      :|||||:|::|||
Db    127  CTGATCATGAGCCGTCACAACCTGCGCCGCGTGGCCAATAACGGCACCAGCCGTGCTGGAA  186

Qy     53  AspValThrProAspAlaTrpProThrTrpProValLysLeuGlyTrpLeuThrProArg  72
           |||||||||||||||||||||:|||||||||||||||| :::
Db    187  CAGTGCGACCAAGGCCTGGCCCAGTGGATGTGCCGGGCGGGCCAGCTGACCACCAGA  246

Qy     73  GlyGlyGluIleAlaTyrLeuGlyHisTyrLeuArgGlnArgLeuValAlaAspGly  92
           ||||||||:|||||||||||::|||    |||||||||||||||:|||:::
Db    247  GGCGGTGTGCTGGAGTGTATATGGGCACTATATGCCGCGAATGGCTGCCGCAGAGAAG  306

Qy     93  LeuLeuAlaAlaLysLysGlyCysProGlnSerGlyGlnValAlaAlaIleIleValAsp  112
           ||||:::           ||||  :||||||||||||||||::||
Db    307  CTGGTGACCAGCCGAGTGCCCGCCCCCGAAAACGCGGTTTATGCCTACGCTAACAGCCTG  366
```

Search completed: June 14, 2006, 15:34:29
Job time : 252.541 secs

```
                GenCore version 5.1.9
        Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:      June 14, 2006, 10:38:44 ; Search time 562.615 seconds
                              (without alignments)
                        8179.104 Million cell updates/sec

Title:       US-10-601-319-2
Perfect score: 2302
Sequence:    1 MKAILIPFLsLLIPLTPQSA..........IVNEARIPACSLRSHHHHHH 440

Scoring table: BLOSUM62
             Xgapop 10.0 , Xgapext   0.5
             Ygapop 10.0 , Ygapext   0.5
             Pgapop  6.0 , Pgapext   7.0
             Delop   6.0 , Delext    7.0

Searched:    5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters:      10489840

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10601319/runat_12062006_180047_27417/app_query.fasta_1
-DB=N Geneseq -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=absss06h
-USER=US10601319 @CGN_1_1_749 @runat_12062006_180047_27417 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :   N Geneseq 8:*
             1: geneseqm1980s:*
             2: geneseqm1990s:*
             3: geneseqm2000s:*
             4: geneseqm2001as:*
             5: geneseqm2001bs:*
             6: geneseqm2002as:*
             7: geneseqm2002bs:*
             8: geneseqm2003as:*
             9: geneseqm2003bs:*
            10: geneseqm2003cs:*
            11: geneseqm2003ds:*
            12: geneseqm2004as:*
            13: geneseqm2004bs:*
            14: geneseqm2005s:*
            15: geneseqm2006s:*

                        SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

Result No.	Score	% Match	Query Length	DB	ID	Description
1	2302	100.0	1323	5	AAC88885	Aac88885 Escherich
2	2302	100.0	1323	6	AAD36473	Aad36473 Escherich
3	2302	100.0	1323	6	AAD25460	Aad25460 Escherich

4	2302	100.0	1323	ADA19445	8	Ada19445 E. coli B
5	2302	100.0	1323	ADO50291	12	Ado50291 Escherich
6	2282	99.1	1323	AEE75412	15	Aee75412 E. coli B
7	2279	99.0	1323	ADO50303	12	Ado50303 Escherich
8	2235	97.1	1323	ADO50301	12	Ado50301 Kangaroo
9	2235	97.1	1299	ABK12514	3	Abk12514 DNA encod
10	2235	97.1	1299	ACA19297	8	Aca19297 Prokaryot
11	2235	97.1	1299	ADI16137	12	Adi16137 Escherich
12	2235	97.1	1901	AAD25463	6	Aad25463 Escherich
13	2235	97.1	1901	ADA19449	8	Ada19449 E. coli K
14	2235	97.1	1901	ADC87744	10	Adc87744 DNA encod
15	2235	97.1	1901	ADO50297	12	Ado50297 Escherich
16	2235	97.1	1901	AED50819	14	Aed50819 Escherich
17	2235	97.1	1901	AEE75418	15	Aee75418 E. coli B
18	2235	97.1	3470	AAC68298	3	Aac68298 R15/APPA
19	2235	97.1	4060	AAC68296	3	Aac68296 R15/APPA
20	2235	97.1	6116	AAC68297	3	Aac68297 R15/APPA
21	2235	97.1	6708	AAC68295	3	Aac68295 R15/APPA
22	2235	97.1	17732	AAC68300	3	Aac68300 Lama2/APP
23	2235	97.1	20623	AAC68294	3	Aac68294 Lama2/APP
24	2230	96.9	1901	ADI16139	12	Adi16139 Shigella
25	2228	96.8	1901	ADO50296	12	Ado50296 Escherich
26	2224	96.6	1296	ADI16138	12	Adi16138 Shigella
27	2221	96.5	1489	AAA28216	3	Aaa28216 E. coli a
28	2221	96.5	1489	ACC57672	8	Acc57672 Escherich
29	2221	96.5	1901	ADO50295	12	Ado50295 Escherich
30	2221	96.5	1901	AEE75417	15	Aee75417 E. coli p
31	2221	96.5	1901	AEE75416	15	Aee75416 E. coli p
32	2218	96.4	1315	ADI02197	10	Adi02197 DNA encod
33	2217	96.3	1486	AAD06831	4	Aad06831 E. coli a
34	2205	95.8	5421	AAC68299	3	Aac68299 SV40/APPA
35	2197	95.4	1486	AAD06832	3	Aad06832 E. coli a
36	2197	95.4	1486	ACC57673	8	Acc57673 Escherich
37	2194.5	95.3	1272	AAX26540	2	Aax26540 DNA encod
38	2188	95.0	1308	ADC87742	10	Adc87742 DNA encod
39	2188	95.0	1308	ADO50299	12	Ado50299 Escherich
40	2188	95.0	1308	AEE75420	15	Aee75420 E. coli B
41	2186	95.0	1299	ADW76362	13	Adw76362 Phytase A
42	2185	94.9	1299	AEF15218	15	Aef15218 Escherich
43	2171	94.3	1296	AED50817	14	Aed50817 Escherich
44	2171	94.3	1901	AEE75422	15	Aee75422 E. coli B
45	2147	93.3	1901	ADA19452	8	Ada19452 E. coli K

```
                         ALIGNMENTS

RESULT 1
AAC88885
ID   AAC88885 standard; DNA; 1323 BP.
XX
AC   AAC88885;
XX
DT   07-MAR-2001  (first entry)
XX
DE   Escherichia coli B phytase enzyme nucleotide sequence.
XX
KW   Escherichia coli B; phytase enzyme; anabolic; phytate digestion;
KW   nutrition; ds.
XX
OS   Escherichia coli.
XX
PN   WO200071728-A1.
XX
PD   30-NOV-2000.
XX
PF   25-MAY-2000; 2000WO-US014846.
XX
PR   25-MAY-1999;  99US-00318528.
XX
PA   (DIVE-) DIVERSA CORP.
XX
PI   Short JM, Kretz KA;
XX
```

DR   WPI; 2001-112081/12.
DR   P-PSDB; AAB37892.
XX
PT   Improving the nutritional value of phytate-containing foodstuffs, using
PT   phytase enzymes which catalyze the liberation of inorganic phosphate from
PT   the phytates.
XX
PS   Claim 2; Fig 1; 147pp; English.
XX
CC   The present sequence encodes a phytase enzyme from Escherichia coli B.
CC   The enzyme catalyses the liberation of inorganic phosphate from the
CC   phytate in phytate-containing foodstuffs and can thus be used to improve
CC   the nutritional value of phytate rich ingredients
XX
SQ   Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:
Score: 3.15e-202
                        2302.00
Percent Similarity: 100.0%   Length: 1323
Best Local Similarity: 100.0%   Matches: 440
Query Match: 100.0%   Conservative: 0
DB: Mismatches: 0
                                      Indels: 0
                                      Gaps: 5

US-10-601-319-2 (1-440) x AAC88885 (1-1323)

Qy   1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db   1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60

Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleIleSerArgHisGly 40
Db   61 TTTGCTCAGAGTGAGCCGGAACTGAAAGTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db   121 GTGCGGGCTCCAACAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy   61 ThrTrpProValLysGlyGlyTrpProThrArgGlyGlyGlyLeuLeuIleAlaTyrLeu 80
Db   181 ACCTGGCCGGTAAAACTGGTTGGCTGGCACCGCGNGGTGGTGAGCTAATCGCCTATCTC 240

Qy   81 GlyHisTyrGlnArgGlnAsnArgLeuValAlaAspProLeuLeuLeuAlaLysGlyCysGlyCysPro 100
Db   241 GGACATTACCAGCGCCAGCCGCGTCTGGTAGCCGACCGGATTGCTGGCGAAAAAAGGGGCTGCCCG 300

Qy   101 GlnSerGlyGlnValAlaIleIleAlaAlaAspProValAspAspThrGlyArgLysThrArgLysThrGlu 120
Db   301 CAGTCTGGTCAGGTGGCGATTATTGCTGCAGACCCGTACCAGCGTAAAACAGGCGAA 360

Qy   121 AlaPheAlaAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db   361 GCCTTCGCCGCCGGCGGTGGCGACCGTGCAATAACCGTACATACCCAGGCAGACACG 420

Qy   141 SerSerProAspProProPheAsnProLeuLysThrGlyValCysGlnLeuGluAlaAsnAla 160
Db   421 TCCAGTCCCGATCCGTTATTAATCCTCTAAAAACTGGCGTTGCCAACTGGATAACGCG 480

Qy   161 AsnValThrAspAlaAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db   481 AACGGTGACTGACGCGATCCTCAGCAGGGCAGGAGCGGTCAATTGCTGACTTTACCGGGCAT 540

Qy   181 ArgGlnThrArgAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerPheAsnLeuCys 200
Db   541 CGGCAAACGGCGTTCCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600

Qy   201 LeuLysLysGlnGlnLysAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db   601 CTTAAACGTGGAGAAACAGGAGAAAGCTGTGTTCATTAACGCAGCAGCATTACCATCGGAACTC 660

Qy   221 LysValSerArgAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db   661 AAGGTGAGCCGCGCAGACAATGTCTCATTAACCGGTGCGGTAAGCCTGGCTAAGTGACG 720

Qy   241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrProGlyArgIleThr 260
Db   721 GAGATATTTCTCCTGCACAAGCACAGGAGATGCCGGAGCCGGGGTGGGAGGGGATCACC 780

Qy   261 AspSerHisGlnTrpAsnThrLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db   781 GATTCACCACCAGTGGAACACCTTGCATAAGCGCCAATTTTATTTGCTACAA 840

Qy   281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db   841 CGCACGCCAGAGTTGCCCGCAGCCGGCCACCCGTTATTGGATTTGATCATGGCAGCG 900

Qy   301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db   901 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGACATTACCCCACTTCAGTACTG 960

Qy   321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db   961 TTTATTGCCGGACACGATACTAATCTGCAAATCTGGGCACTGGAGCTCAACTGG 1020

Qy   341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1021 ACGCTTCCCGGTCAGCCCGATAACACGCCGCCGATGGTGGTGAACTGGTTTGAACGCTGG 1080

Qy   361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1081 CGTCGGCTAAGGGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTTCCAGACTTTACAG 1140

Qy   381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyValValLysLeuThr 400
Db 1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCCGGAGAGGTGAAACTGACC 1200

Qy   401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGGCAGGATGTGAAGAGCGAAATGCCAGGGCATGTGTCGTTGGCAGGTTTTACGCAA 1260

Qy   421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHisHis 440
Db 1261 ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTGAGATCTCATCACCATCAC 1320

```
DR   WPI; 2002-164149/21.
DR   P-PSDB; AAE22836.
XX
PT   New dietary aids comprising sustained release biocompatible compositions,
PT   comprise agent that assists in digestion, useful for delivering enzymes,
PT   therapeutics, medicine or agents to an organism.
XX
PS   Disclosure; Fig 1; 89pp; English.
XX
CC   The present invention relates to novel dietary aids comprising sustained
CC   release biocompatible composition which comprises an agent (enzymes such
CC   as phytase, amylase, esterase, protease) that assists in digestion. The
CC   biocompatible composition is effective upon oral consumption and release
CC   in the digestive tract of a subject. The dietary aids are useful for
CC   delivering enzymes, therapeutics, medicine and agents to an organism. The
CC   use of enzymes and other agents in digestive aids of livestock or
CC   domesticated animals not only improves the animal's health and life
CC   expectancy but also assists in increasing the health of livestock or in
CC   the production of foodstuffs from livestock. The present sequence is
CC   Escherichia coli phytase DNA
XX
SQ   Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:
Score:                  3.15e-202    Length:          1323
                        2302.00      Matches:         440
Percent Similarity:     100.0%       Conservative:    0
Best Local Similarity:  100.0%       Mismatches:      0
Query Match:            100.0%       Indels:          0
DB:                                  Gaps:            6

US-10-601-319-2  (1-440) x AAD36473 (1-1323)

Qy    1 MetLysAlaIleLeuIleProPheLeuSerLeuIleLeuProLeuThrProGlnSerAla  20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA  60

Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly  40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy   61 ThrTrpProValLysGlyTyrProLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu  80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGGNGAGCTAATCGCCTATCTC 240

Qy   81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyLeuLysPro 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GGACATTACCACCAGTGGATCCGGACGGATTGCTGGCGAAAAAGGGGCTGCCCG        300

Qy  101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA 360

Qy  121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GCCTTTGCCGCCGGCGGTGGCACCTGACCTGCAATAACCGTACATACCCAGGCAGATACG 420

Qy  141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TCCAGTCCCGGATCCGGTTATTTAATCCTCTAAAAACTGGCGTTTGCCACTGGATAACGCG 480

Qy  161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 AACGTGACTGACGCGATCCTCAGCAGGGCAGCAGGAGGTCAATTGCTGACTTTACCGGGCAT 540

Qy  181 ArgGlnThrAlaPheArgGluLeuGluLeuGluValLeuAsnPheProGlnSerAsnLeuCys 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 CGGCAAACGGCGTTTCGCGAACTGGAACTGGAAGTGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600
```

```
Qy  201 LeuLysArgGluLysGlnAlaGluGlySerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CTTAAACGTGAGAAACAGGACGAAGCTGTTCATTAACCGAGGCATTACCATCGGAACTC 660

Qy  221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 AAGGTGAGCCGCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 720

Qy  241 GluIlePheLeuLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyValArgIleThr 260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GAGATATTTTCTCTGCAACAAGCACAGGGAATGCCCGAGCCGGGGTGGGAAGGATCACC 780

Qy  261 AspSerHisGlnTrpProAsnThrLeuLeuSerLeuHisAsnAlaGlnProPheTyrLeuLeuGln 280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GATTCACCACCAGTGGAACACCTTGCTAAGTTGCATAACGCGCAATTTTATTTGCTACAA 840

Qy  281 ArgThrProGluValAlaAlaSerArgArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTGGATTTGATCATGGCAGCG 900

Qy  301 LeuThrProHisProProGlnLysGlnAlaTyrGlnTyrGlyTyrValThrLeuProThrSerValLeu 320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTACTG 960

Qy  321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG 1020

Qy  341 ThrLeuProGlyGlnProLeuAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 ACGCTTCCCGGTCAGCCGCTGAATACACGCCCCAGGTGGTGAACGGTGTTTGAACGCTGG 1080

Qy  361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnIleThrLeuGln 380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 CGTCGGCTAAGCGATAATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTTCAGACTTTACAG 1140

Qy  381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCCGGAGGAGGTGAAACTGACC 1200

Qy  401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA 1260

Qy  421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHisHis 440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 ATCGTGAATGAAGCACGCATACCGGGCGTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320
```

```
RESULT 3
AAD25460
ID   AAD25460 standard; DNA; 1323 BP.
XX
AC   AAD25460;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Escherichia coli B phytase DNA
XX
KW   Bacterial phytase; K12 appA phytase; protease stability; anabolic;
KW   gastrointestinal; nutritional value; feed treatment process; therapy;
KW   thermal tolerance; growth performance; alcoholic drink; biopulping;
KW   non-alcoholic drink; bioleaching; B phytase; ds.
XX
OS   Escherichia coli.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..1323
FT                  /tag= a
FT                  /product= "E. coli B phytase protein"
XX
PN   WO200190333-A2.
XX
PD   29-NOV-2001.
```

```
Db  361 GCCTTCGCCGCCGGCGGCTGGCACCTGACTGTGCAATAACCGTACATAACCCAGGCAGATACG 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  141 SerSerProAspProLeuPheAsnProLeuLysThrGlnProLeuLysGlnLeuAspAsnAla 160
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TCCAGTCCGATCCGTATTTAATCCTTAAAAACGGGCGTTTGCCAACTGGATAAACCGCG 480
Qy  161 AsnValThrAspAlaLeuLeuSerArgAlaGlyLeuIleAlaAspPheThrGlyHis 180
Db  481 AACGTGACTGACGCGATCCTCAGCAGGCAGGAGGGTCAATTGCTGACTTACCGGGCAT 540
Qy  181 ArgGlnThrAlaPheArgGluLeuGluLeuLeuArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db  541 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600
Qy  201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db  601 CTTAAACGTGAGAAACACAGGACGAAAGCTGTTCATTAACGACGCATTACCATCCGAACTC 660
Qy  221 LysValSerAlaAspAspAsnValSerLeuThrGlyAlaValSerLeuValSerMetLeuThr 240
Db  661 AAGGTGAGCGCCACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 720
Qy  241 GluIlePheLeuLeuGlnIleAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db  721 GAGATATTTCTCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGTGGGGAAGGATCACC 780
Qy  261 AspSerHisGlnTrpAsnThrLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db  781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCCCAATTTTATTTGCTACA 840
Qy  281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuMetAlaAla 300
Db  841 CGCACGCCAGAGGTTGCCCGCCAGCCCGCGCCACCCCGTTATTGATTTGATCATGGCAGCG 900
Qy  301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db  901 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGGACATTACCCACTTCAGTACTG 960
Qy  321 PheIleAlaGlyHisAspPheThrAsnLeuAlaAsnLeuGlyAlaLeuGlyLeuAsnTrp 340
Db  961 TTTATTGCCGGACACGATACTAATCTGGCAAATCTCCGGCGGCCACTGGAGCTCAACTGG 1020
Qy  341 ThrLeuProGlyGlnThrProAspAsnThrProGlyGlyLeuValPheGluValArgTrp 360
Db 1021 ACGCTTCCCGGTCAGCCGGGATAACACGCCGCCAGGTGGTGGAACTGGTGTTGAACGCTGG 1080
Qy  361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGlnIle 380
Db 1081 CGTCGGCTAAGGCGGATAACAGCCAGTGGATTCGCTGGTCTTCCAGACTTTACAG 1140
Qy  381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProGlyGlyLeuValLysLeuThr 400
Db 1141 CAGATGCGTGATAAAACGCCGTGTCATTAAATAAATACGCCCGGAGAGGTGAAACTGACC 1200
Qy  401 LeuAlaGlyCysGluLeuArgAsnAlaGlnAlaGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGGCAGGATGTGAAGAGCGAAATGCCAGGGGCATGTGTTCGGTGGCAGGTTTTACGGCA 1260
Qy  421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHisHis 440
Db 1261 ATCGTGAATGAAGCAGCACCGCCGTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320
```

CC   The patent discloses recombinant bacterial phytase from Escherichia coli
CC   K12 appA phytase. The enzyme has phytase activity and improved thermal
CC   tolerance when compared with wild-type phytase. It has improved protease
CC   stability at low pH. The recombinant phytase is useful for improving the
CC   nutritional value of phytate-containing foodstuffs and subsequently
CC   improving the growth performance of an organism that consumes it, in
CC   treating animal digestive systems, in feed treatment processes and for in
CC   vitro purposes related to research, discovery and development. They are
CC   also used for generating recombinant digestive system life forms, for
CC   producing or manufacturing alcoholic and non-alcoholic drinks based on
CC   the use of moulds, grains and/or plants, in bioplping and bio- bleaching
CC   where a reduction in the use of environmentally harmful chemicals that
CC   are traditionally used in the pulp and paper industry is desired and in
CC   the reduction or possible elimination of the need for mineral
CC   supplements, enzymes or therapeutic drugs for animals from the daily feed
CC   thus increasing the amount calories and nutrients present in the feed.
CC   The present sequence is a DNA encoding E. coli B phytase protein
XX
SQ   Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;

```
Alignment Scores:
Pred. No.:            3.15e-202     Length:         1323
Score:               2302.00       Matches:        440
Percent Similarity:  100.0%        Conservative:   0
Best Local Similarity: 100.0%      Mismatches:     0
Query Match:         100.0%        Indels:         0
DB:                                Gaps:           6
```

US-10-601-319-2 (1-440) x AAD25460 (1-1323)

```
Qy    1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTGCTTATTCCGTTAACCCCGCAATCTGCA 60
Qy   21 PheAlaGlnSerGlyLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db   61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGGTGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120
Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db  121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180
Qy   61 ThrTrpProValLysGluGlyTyrLeuThrProArgGlyGlyGluLeuIleLeuAlaTyrLeu 80
Db  181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC 240
Qy   81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuLeuAlaLysGlyCysPro 100
Db  241 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGCTGCGGAAAAAGGGCTGCCCG 300
Qy  101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db  301 CAGTCTGGTCAGGTGGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA 360
Qy  121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
```

```
KW   Phytase; ds; gene; phytate; appA gene; animal feed; inorganic phosphate;
KW   digestion enhancement; transgenic; thermal tolerance; protease stability.
XX
OS   Escherichia coli B.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..1323
FT                  /*tag= a
FT                  /product= "Phytase"
XX
PN   US2002136754-A1.
XX
PD   26-SEP-2002.
XX
PF   24-MAY-2001; 2001US-00866379.
XX
PR   13-AUG-1997; 97US-00910798.
PR   01-MAR-1999; 99US-00259214.
PR   13-APR-1999; 99US-00291931.
PR   25-MAY-1999; 99US-00318528.
PR   25-MAY-2000; 2000US-00580515.
XX
PA   (SHOR/) SHORT J M.
PA   (KRET/) KRETZ K A.
PA   (GRAY/) GRAY K A.
PA   (BART/) BARTON N R.
PA   (GARR/) GARRETT J B.
PA   (DONO/) O' DONOGHUE E.
PA   (MATH/) MATHUR E J.
XX
PI   Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB;
PI   O' Donoghue E, Mathur EJ;
XX
DR   WPI; 2003-040002/03.
DR   P-PSDB; ADA19446.
XX
PT   Isolated Escherichia coli polynucleotide encoding a modified phytase
PT   enzyme, useful in the production of animal feed, for improving the
PT   nutritional value of phytate-containing foodstuff and for enhancing
PT   digestion in humans and animals.
XX
PS   Claim 2; Fig 1; 62pp; English.
XX
CC   The invention relates to an isolated Escherichia coli polynucleotide
CC   encoding a phytase enzyme appearing as ADA19450 and having amino acids
CC   modification from W68B, Q84W, A95P, K97C, S168B, R180Y, N226C or Y277D.
CC   Also included the E. coli appA gene ADA19449 (or an oligonucleotide
CC   derived from it) or its mutant sequence ADA19452, expression vectors,
CC   host cells, a method of improving nutritional value of a phytate-
CC   containing foodstuff by contacting the phytate-containing foodstuff with
CC   a pure phytase enzyme mutated phytase (where the phytase enzyme catalyzes
CC   the liberation of inorganic phosphate from the phytate in the phytate-
CC   containing foodstuff), a method to produce an animal feed containing a
CC   microbial phytase (comprising culturing the plant cell, plant part or
CC   plant under conditions where the nucleotide sequence is expressed and
CC   converting the plant cells, plants or plant into a composition for animal
CC   feed), a feed composition for animals (comprising the plant seeds, plant
CC   cells, plant parts or plants in admixture with a phytate-containing
CC   foodstuff), a method to treat a human or an animal able to benefit from
CC   digestive enhancement by the activity of an exogenous phytase enzymes
CC   comprising administering to the human or animal the plant seed, plant
CC   cells, plant parts or plants of a transgenic plant which is modified to
CC   contain an expression system which expresses a nucleotide sequence
CC   encoding a phytase enzyme, a transgenic non-human organism whose genome
CC   comprising a heterologous nucleic acid sequence encoding a polypeptide
CC   having phytase activity. The phytase enzyme is useful for improving the
CC   nutritional value of phytate-containing foodstuff, in the production of
CC   animal feed and for enhancing digestion in humans and animals. The
CC   invented method improves thermal tolerance and protease stability. It
CC   also improves the feeding value of phytate rich ingredients. The present
CC   sequence encodes E. coli B wild-type phytase.
XX
SQ   Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;
```

```
Alignment Scores:
Pred. No.:                       1323
Score:              3.15e-202    Matches:         2302.00
Percent Similarity: 100.0%       Conservative:       0
Best Local Similarity: 100.0%    Mismatches:         0
Query Match:        100.0%       Indels:             0
DB:                              Gaps:               8

US-10-601-319-2 (1-440) x ADA19445 (1-1323)

Qy     1 MetLysAlaIleLeuIleProPhLeuSerLeuIleProLeuThrProGlnSerAla 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAAAGCGATCTTAATCCATTTTATCTCTTCTGATTCCGTTAACCCGGCAAATCTGCA 60

Qy    21 PheAlaGlnSerGluProGluLeuLysGluSerValValIleValSerArgHisGly 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TTCGCTCAGAGTGAGCCGGAGCTGAAGTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120

Qy    41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAsnValThrProAspAlaTrpPro 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy    61 ThrTrpProValLysLeuGlyTrpLeuThrProLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCNGGTGGTGACTAATCGCCTATCTC 240

Qy    81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyLysCysPro 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGACATTACCAACGCCAGGGTCTCGGTAGCCGACGATTGCTGGCAGAAAAGGGCTGCCCG 300

Qy   101 GlnSerGlyGlnValAlaIleIleAlaAspValAspAlaArgThrArgLysThrGlyGlu 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGTCTGGTCAGGTCGCGGATTATTGCTGCGATGTCGACGAGCGTACCCGTAAAACAGGCGAA 360

Qy   121 AlaPheAlaAlaGlyLeuValAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCCTTCGCCGCCGGCCTGCCACCTGACCTGCATAACCGTACATACCCAGGCAGATACG 420

Qy   141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCCAGTCCGGATCCGTTATTTAATCCTCTAAAACTGGGTTGCCAACTGGATAACGCG 480

Qy   161 AsnValThrAspAlaIleLeuSerArgAlaAlaGlyGlySerIleAlaAspPheThrGlyHis 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AACGTGACTGACGCGATCCTCAGCAGGGCAGCAGGAGGGTCAATGCTGGACTTTACCGGCCAT 540

Qy   181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC 600

Qy   201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAAGTC 660

Qy   221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 720

Qy   241 GluIlePheLeuLeuGlnIleAlaGlnAspGlyMetProGluProGlyTrpGlyArgIleThr 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GAGATATTTCTCCTGCAACAAGCACAGGATGGAATGCCGGAGCCGGGGTGGGGAAGGATCACC 780

Qy   261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnInPheTyrLeuLeuGln 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GATTCACACCAGTGGAACACCTTGCTAAGTTGCATAACGCGCCAATTTTATTTGCTACAA 840

Qy   281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTGGATTTGATCATGGCAGCG 900

Qy   301 LeuThrProHisProGlnProGlnInAlaTyrGlyGlnValThrLeuProThrSerValLeu 320
```

```
Db    901 TTGACGGCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTACTG  960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp     340
Db    961 TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluValPheGluArgTrp     360
Db   1021 ACGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGTGGTGAACTGGTGTTTGAACCGTGG 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln  380
Db   1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr  400
Db   1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCAGGAGGTGAAACTGACC   1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    401 LeuAlaGlyCysGluGluArgAsnAlaGlnLeuMetCysSerSerLeuAlaGlyPheThrGln 420
Db   1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGCATGTGTTCGTTGGCAGGTTTACGCAA  1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHisHis  440
Db   1261 ATCGTGAATGAAGCACGCATACCGGCTGCAGTTTGAGATCTCATCACCACCATCAC    1320
```

RESULT 5
AD050291
ID   AD050291 standard; DNA; 1323 BP.
XX
AC   AD050291;
XX
DT   29-JUL-2004 (first entry)
XX
DE   Escherichia coli B phytase DNA.
XX
KW   Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW   phytate; animal feed; fish feed; dough; baking; gene; ds.
XX
OS   Escherichia coli; B.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1323
FT   /tag= a
FT   /product= "Phytase enzyme"
FT   unsure   214..216
FT   /tag= b
FT   /note= "Encodes Arg"
XX
PN   US2004091968-A1.
XX
PD   13-MAY-2004.
XX
PF   20-JUN-2003; 2003US-00601319.
XX
PR   13-AUG-1997;   97US-00910798.
PR   01-MAR-1999;   99US-00259214.
PR   13-APR-1999;   99US-00291931.
PR   25-MAY-1999;   99US-00318528.
PR   25-MAY-2000;   2000US-00580515.
PR   24-MAY-2001;   2001US-00866379.
XX
PA   (SHOR/) SHORT J M.
PA   (KRET/) KRETZ K.
PA   (GRAY/) GRAY K A.
PA   (BART/) BARTON N R.
PA   (GARR/) GARRETT J B.
PA   (ODON/) O'DONOGHUE E.
PA   (MATH/) MATHER E J.
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI   Mather EJ;
XX

DR   WPI; 2004-374952/35.
DR   P-PSDB; AD050292.
XX
PT   Producing phytase, involves providing nucleic acid derived from bacteria
PT   encoding polypeptide having phytase activity, and expressing nucleic acid
PT   in yeast.
XX
PS   Claim 3; SEQ ID NO 1; 74pp; English.
XX
CC   The invention relates to a method for producing phytase that involves
CC   providing a nucleic acid encoding phytase derived from a bacteria, and
CC   expressing the nucleic acid in a yeast under conditions that allow
CC   expression of the enzyme in the yeast. The invention also relates to
CC   modified phytase enzyme which has improved thermal tolerance and protease
CC   stability at low pH. The phytase enzyme can be used in foodstuffs to
CC   improve the feeding value of phytate rich ingredients, and in diet of
CC   numerous animals including mammals, fowls and fishes, commercially
CC   significant mammals such as pigs, goats, laboratory rodents, commercially
CC   significant avian species such as chicken, ducks, doves, parrot, etc.,
CC   commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC   etc., in dough making and baking, in dietary aids for animals. The method
CC   provides easy manufacture of the active ingredient loaded biocompatible
CC   composition, higher yields and loading efficiency. The phytase
CC   incorporated in to the dietary aid is safe for animals. The present
CC   sequence is Escherichia coli B phytase DNA.
XX
SQ   Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;
XX

Alignment Scores:
Pred. No.:   Length: 1323
Score: 2302.00   Matches: 440
Percent Similarity: 100.0%   Conservative: 0
Best Local Similarity: 100.0%   Mismatches: 0
Query Match: 100.0%   Indels: 0
DB: 12   Gaps: 0

US-10-601-319-2 (1-440) x AD050291 (1-1323)

Qy   1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db   1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCGAATTCCGTTAACCCCGCAATCTGCA 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db   61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro   60
Db   121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   61 ThrTrpProValLysLeuGlyTrpLeuThrProLeuTrpArgGlyGlyLeuLeuIleAlaTyrLeu 80
Db   181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCNGGTGGTGAGCTAATCGCCTATCTC 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   81 GlyHisTyrGlnArgGlnValAlaAspArgLeuValAlaAspLysLysGlyCysPro   100
Db   241 GGACATTACCAACGCCAGGTCGTCGGACGAGGATTGCTGGCAGAAAAAGGGCTGCCCG   300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   101 GlnSerGlyGlnValAlaIleIleAlaIleAsnSerAla   120
Db   301 CAGTCTGGTCAGTCGCGGATTATTGCTATTAATCCTAAAAACTGGCGTTTGCCAACTGGA 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   121 AlaPheAlaAlaGlyLeuAlaProAspCysSerAlaIleThrValHisThrGlnAlaAspThr 140
Db   361 GCCTTCGCGCCGGGCTGGCACCTGACCTGTGCAATAACCGTACATACCCAGGCAGATACG 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   141 SerSerProAspPheLeuAsnProLeuLysThrGlyValCysGlnLeuAspLeuAsnAla 160
Db   421 TCCAGTCCGGATCCGTTTTTAATCCTTAAAAACTGGCGTTTGCCAACTGGATAACGCG   480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   161 AsnValThrAspAlaAlaLeuSerArgAlaGlySerIleAlaAspPheGlyThrGlyHis 180
Db   481 AACGTGACTGACGCGGCATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTACGGGGCAT 540
```

```
Qy    181  ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProSerAsnLeuCys  200
                                                                       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  CGGCAAACGGCGTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTGTGC  600

Qy    201  LeuLysArgGluLysGlnSerGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220
                                                                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  CTTAAACGTGAGAAACAGAGCGAAAGCTGTTCATTAACGACAGGCATTACCATCGGAACTC  660

Qy    221  LysValSerAlaAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr  240
                                                                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCCGCCAATCAATGCTGACG  720

Qy    241  GluIlePheLeuLeuGlnGlnAlaLeuGlnGlyMetProGluProGlyGlyTrpGlyArgIleThr  260
                                                                       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  GAGATATTTTCTCTGCAACAAGCACAGGAATGCCGGAACCGCCGGGGTGGGGAAGGATCACC  780

Qy    261  AspSerHisGlnThrLeuSerLeuHisAsnAlaAlaGlnPheTyrLeuLeuGln  280
                                                                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  GATTCACACCAGTGGAACACCTTGTCTAAGTTTGCATAACGGCAATTTATTTGCTACAA  840

Qy    281  ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla  300
                                                                       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  CGCACGCCAGAGGTGCCCGCCAGCCCGCCACCCGTTATTGGATTTGATCATGGCAGCG  900

Qy    301  LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu  320
                                                                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTACTG  960

Qy    321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340
                                                                       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCCACTGAGCTCAACTGG  1020

Qy    341  ThrLeuProGlyGlnProAspAsnProProGlyGlyLeuValPheGluValPheArgTrp  360
                                                                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 ACGCTTCCCGGTCAGCCCGATAACCGCCGGGGTGTGGTGAACGTGTTTGAACGCTGG  1080

Qy    361  ArgArgLeuSerAspAsnSerGlnIleTrpIleGlnIleValSerLeuValPheGlnLeuThrLeuGln  380
                                                                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CGTCGGCTAAGCGATAACAGCCACAGTCCACGTTCCGCTGGATTCAGGTTCCGCAGACTTTACAG  1140

Qy    381  GlnMetArgAspLysThrProLeuGluSerLeuAsnThrProProGlyGluValLysLysLeuThr  400
                                                                       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 CAGATGCGTGATAAAACCCCGCTGTCATTAAATACGCCGCCCGGAGAGGTGAAACTGACC  1200

Qy    401  LeuAlaGlyLysCysGluGluArgAsnAlaGlnLysGlyMetCysSerLeuAlaGlyPheThrGln  420
                                                                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGCATGTGTTCGTTGGGCAGGTTTTTACGGCA  1260

Qy    421  IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHis  440
                                                                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTGAGATCTCATCACCATCAC  1320
```

```
RESULT 6
AEE75412
ID   AEE75412 standard; DNA; 1323 BP.
XX
AC   AEE75412;
XX
DT   23-FEB-2006  (first entry)
XX
DE   E. coli B phytase coding sequence SEQ ID NO: 1.
XX
KW   ds; coding sequence; thermostable; phytase; protein stabilization;
KW   pharmaceutical; fertilizer; cramp; cell culture; osteoporosis;
KW   osteopathic; food.
XX
OS   Escherichia coli B.
XX
FH   Key          Location/Qualifiers
FT   CDS          1..1323
FT                /*tag= a
FT                /product= "Phytase"
```

```
FT   unsure       214..216
FT                /*tag= b
FT                /note= "ecodes Arg"
XX
PN   US2005281792-A1.
XX
PD   22-DEC-2005.
XX
PF   01-SEP-2004; 2004US-00933115.
XX
PR   13-AUG-1997;    97US-00910798.
PR   01-MAR-1999;    99US-00259214.
PR   13-APR-1999;    99US-00291931.
PR   25-MAY-1999;    99US-00318528.
PR   25-MAY-2000;   2000US-00580515.
PR   24-MAY-2001;   2001US-00866379.
XX
PA   (SHOR/)  SHORT J M.
PA   (KRET/)  KRETZ K A.
PA   (GRAY/)  GRAY K A.
PA   (BART/)  BARTON N R.
PA   (GARR/)  GARRETT J B.
PA   (ODON/)  ODONOGHUE E.
PA   (BAUM/)  BAUM W.
PA   (ROBE/)  ROBERTSON D E.
PA   (ZORN/)  ZORNER P.
XX
PI   Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, Odonoghue E;
PI   Baum W, Robertson DE, Zorner P;
XX
DR   WPI; 2006-055022/06.
DR   P-PSDB; AEE75413.
XX
PT   Formulation useful as dietary supplement for treating, preventing or
PT   reversing osteoporosis and bone loss, and preventing muscle cramps,
PT   comprises one or more polypeptides having phytase activity.
XX
PS   Claim 1; SEQ ID NO 1; 82pp; English.
XX
CC   The present sequence is that of a polynucleotide encoding the Escherichia
CC   coli B phytase. The present invention relates to a novel formulation,
CC   comprising at least one polypeptide having phytase activity, where the
CC   polypeptide is a modified sequence (AEE75421) derived from the
CC   Escherichia coli K-12 appA phytase (AEE75419). The modification of the
CC   enzyme, comprising a series of mutations, improves the thermal tolerance
CC   and protease stability of the protein. The specification also claims a
CC   pharmaceutical composition; a kit; an immobilized phytase; a fertilizer
CC   or soil additive; a liquid supplement for preventing muscle cramps; a
CC   hydrating agent; a tissue culture or cell culture media; and a plant food
CC   additive, all comprising the novel phytase, and a method of reducing
CC   pollution and increasing nutrient availability in an environment or
CC   environmental sample by degrading environmental phytic acid. The novel
CC   enzyme is a dietary supplement useful for treating, preventing or
CC   reversing osteoporosis or bone loss, and preventing muscle cramps. The
CC   liquid supplement is useful for preventing muscle cramps. The method is
CC   useful for reducing pollution and increasing nutrient availability in an
CC   environment or environmental sample by degrading environmental phytic
CC   acid, where the environment or environmental sample comprises a soil or a
CC   body of water. The immobilized phytase is useful in foodstuffs for
CC   improving the feeding value of phytate rich ingredients.
XX
SQ   Sequence 1323 BP; 323 A; 353 G; 357 C; 289 T; 0 U; 1 Other;
XX

Alignment Scores:
Pred. No.:                            Length:         1323
Score:                  3.15e-202     Matches:         440
Percent Similarity:     2302.00       Conservative:      0
Best Local Similarity:  100.0%        Mismatches:        0
Query Match:            100.0%        Indels:            0
DB:                     15            Gaps:              0

US-10-601-319-2 (1-440) x AEE75412 (1-1323)
```

```
Qy   1081  CGTCGGCTAAGCCATACAGCCAGTGGATTCAGTTTCGCTGGTCTTCCAGACTTACAG  1140
Db
Qy    381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr  400
Db   1141  CAGATGCGGTGATAAAACGCCGTCATTAAATACGCCGCCCGGAGAGTGAAACTGACC  1200
Qy    401  LeuAlaGlyCysGluGluArgAsnAlaGlnIleGlyMetCysSerLeuAlaGlyPheThrGln  420
Db   1201  CTTGGCAGGATGTGAAGAGCGAAATGCCAGGGCCATGTGTTCGTTGGCAGTTTTACGCAA  1260
Qy    421  IleValAsnGluAlaGlnIleProAlaCysSerLeuArgSerHisArgSerHisHisHis  440
Db   1261  ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTGAGATCTCATCACCATCACCATAC  1320
```

RESULT 7
ADO50303
```
ID   ADO50303   standard; DNA; 1323 BP.
XX
AC   ADO50303;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Escherichia coli W phytase 875PH2 mutant DNA.
XX
KW   Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW   phytate; animal feed; fish feed; dough; baking; mutant; gene; ds.
XX
OS   Escherichia coli.
XX
FH   Key              Location/Qualifiers
FT   CDS              1..1323
FT                    /product= "Mutant phytase enzyme"
FT                    /*tag= a
FT   unsure           214..216
FT                    /*tag= b
FT                    /note= "Encodes Arg"
FT   mutation         replace(478,G)
FT                    /*tag= c
FT   mutation         replace(527,A)
FT                    /*tag= d
FT   mutation         replace(893,T)
FT                    /*tag= e
FT   mutation         replace(895,G)
FT                    /*tag= f
XX
PN   US2004091968-A1.
XX
PD   13-MAY-2004.
XX
PF   20-JUN-2003; 2003US-00601319.
XX
PR   13-AUG-1997; 97US-00910798;
PR   01-MAR-1999; 99US-00259214;
PR   13-APR-1999; 99US-00291931;
PR   25-MAY-1999; 99US-00318528.
PR   25-MAY-2000; 2000US-00580515.
PR   24-MAY-2001; 2001US-00866379.
XX
PA   (SHOR/)  SHORT J M.
PA   (KRET/)  KRETZ K.
PA   (GRAY/)  GRAY K A.
PA   (BART/)  BARTON N R.
PA   (GARR/)  GARRETT J B.
PA   (ODON/)  O'DONOGHUE E.
PA   (MATH/)  MATHER E J.
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI   Mather EJ;
XX
DR   WPI; 2004-374952/35.
DR   P-PSDB; ADO50292, ADO50304.
XX
```

```
Qy    1  MetLysAlaIleLeuIlePheLeuSerLeuLeuIleProLeuThrProLeuSerAla   20
Db    1  ATGAAAGCGATCTAATCCCATTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA   60

Qy   21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleSerValSerArgHisGly   40
Db   61  TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGGAAAGTGTGGTGATTGTCAGTGTCATGGT  120

Qy   41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro   60
Db  121  GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA  180

Qy   61  ThrTrpProValLysGlyLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu   80
Db  181  ACCTGGCCGGTAAAAGGTTTGGGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC  240

Qy   81  GlyHisTyrGlnArgGlnArgGlyValAlaAspGlyLeuLeuAlaLysGlyLysCysPro  100
Db  241  GGACATTACCACCGCCCAGCGTCTGGTAGCCGACGGATTGCTGGCGAAAAGGGCTGCCCG  300

Qy  101  GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu  120
Db  301  CAGTCTGGTCAGGTGGCGATTATTGCTGACGTCGACGAGCGTACCCGTAAAACAGGCGAA  360

Qy  121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr  140
Db  361  GCCTTCGCCGCCGGGCTGGCACCTGACTGCGCAATAACCGTACATACCCAGGCAGATACG  420

Qy  141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla  160
Db  421  TCCAGTCCCGGATCCGTTATTAATCCTCTAAAAACTGGCGTTGCCAACTGATAACGCG  480

Qy  161  AsnValThrAspAlaLeuLeuSerArgArgAlaGlyGlySerIleLeuAlaAspPheThrGlyHis  180
Db  481  AACGTGACTGACGCGCTCTCCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT  540

Qy  181  ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys  200
Db  541  CGGGCAAACGGCGTTTGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTGTGC  600

Qy  201  LeuLysArgGluGluLysGlnAspCysSerSerLeuThrGlnAlaLeuProSerGluLeu  220
Db  601  CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC  660

Qy  221  LysValSerAlaAspAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr  240
Db  661  AAGGTGAGCGCCGACAATGTCTCATTAACCGGTACGCCTGCGGTCAAATGCTGACG  720

Qy  241  GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr  260
Db  721  GAGATATTTCTCCTGCAACAGAGCAGCAGGGAATGCCCGGAGCGCCGGGGGTGGGGGAAGGATCACC  780

Qy  261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln  280
Db  781  GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATTAACGCGGCCAATTTTATTTGCTACAA  840

Qy  281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla  300
Db  841  CGGCACGCCAGAGGTTGCCCGCAGCCGCGCGCACCCCGTTATTGGATTTGATCATGGCAGCG  900

Qy  301  LeuThrProHisProGlyGlnLysGlyGlnValThrGlyValThrLeuProThrSerValLeu  320
Db  901  TTGACGCCCCATCCAGGCCAAAAACAGGCTATGGTGTGACATTACCCACTTCAGTACTG  960

Qy  321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340
Db  961  TTTATTGCCGGACACGATACTAATCTCGGCGGCGCACTGGAGCTCAACTGG  1020

Qy  341  ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp  360
Db 1021  ACGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG  1080

Qy  361  ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln  380
```

```
PT  Producing phytase, involves providing nucleic acid derived from bacteria
PT  encoding polypeptide having phytase activity, and expressing nucleic acid
PT  in yeast.
XX
PS  Disclosure; Page; 74pp; English.
XX
CC  The invention relates to a method for producing phytase that involves
CC  providing a nucleic acid encoding phytase derived from a bacteria, and
CC  expressing the nucleic acid in a yeast under conditions that allow
CC  expression of the enzyme in the yeast. The invention also relates to
CC  modified phytase enzyme which has improved thermal tolerance and protease
CC  stability at low pH. The phytase enzyme can be used in foodstuffs to
CC  improve the feeding value of phytate rich ingredients, and in diet of
CC  numerous animals including mammals, fowls and fishes, commercially
CC  significant mammals such as pigs, goats, laboratory rodents, commercially
CC  significant avian species such as chicken, ducks, doves, parrot, etc.,
CC  commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC  etc., in dough making and baking, in dietary aids for animals. The method
CC  provides easy manufacture of the active ingredient loaded biocompatible
CC  composition, higher yields and loading efficiency. The phytase
CC  incorporated in to the dietary aid is safe for animals. The present
CC  sequence is Escherichia coli W phytase mutant DNA. Note: The present
CC  sequence is not shown in the specification but has been derived from
CC  Escherichia coli B phytase DNA ADO50291.
XX
SQ  Sequence 1323 BP; 324 A; 353 C; 356 G; 289 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:
Score:                    2282.00
Percent Similarity:         99.3%
Best Local Similarity:      99.1%
Query Match:                99.1%
DB:                           12

US-10-601-319-2 (1-440) x ADO50303 (1-1323)

Qy    1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla  20
Db    1 ATGAAAGCGATCTTAATCCCATTTTATCTCTCTGATTCCGTTAACCCCGCAATCTGCA    60

Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly  40
Db   61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT  120

Qy   41 ValArgAlaProThrLysAlaThrProLeuMetGlnAsnValThrProAspAlaTrpPro  60
Db  121 GTGCGTGCTCCAACCAAGGCCACGCCACTGATGCAGAATGTCACCCCAGACGCATGGCCA  180

Qy   61 ThrTrpProValLysGlyLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu  80
Db  181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCCGACGGNGGTGGNGAGCTAATCGCCTATCTC  240

Qy   81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro  100
Db  241 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG  300

Qy  101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu  120
Db  301 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA  360

Qy  121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr  140
Db  361 GCCTTCGCCGCGGGTCTGGCACCTGACTGCGCAATAACCGTACATACCCAGGCAGATACG  420

Qy  141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla  160
Db  421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAACTGGCGTTTGCCAACTGGATAACTCG  480

Qy  161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis  180
Db  481 AACGTGACTGACGCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGGCTTTACCGGGCAT  540

Qy  181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys  200
Db  541 CGGCAAACGGCGTTTCGCGAACTGGAACGTGTTCTAAATTTCCCGCAATCAAACTTGTGC  600

Qy  201 LeuLysArgGluLysGlnGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220
Db  601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC  660

Qy  221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr  240
Db  661 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTTCGCATCAATGCTGACG  720

Qy  241 GluIlePheLeuLeuGlnGlnAlaGlnAsnLeuAlaIleAsnGlyMetProGluProGlyTrpGlyIleThr  260
Db  721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGTGGGGGAAGGATCACC  780

Qy  261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln  280
Db  781 GATTCACACCAGTGGAACACTTTGCTAAGTTGCATAACCGCAATTTTATTGCTACAA  840

Qy  281 ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla  300
Db  841 CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTGGATTTGATCAAGACAGCG  900

Qy  301 LeuThrProHisProProGlnLysGlnLysAlaTyrGlyValThrLeuProThrSerValLeu  320
Db  901 TTGACGCCCCATCCACCGCAAAAACAGGCTATGGTGACATTACCCACCTTCAGTACTG  960

Qy  321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340
Db  961 TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG  1020

Qy  341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp  360
Db 1021 ACGCTTCCCGGTCAGCCCGATAACACGCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG  1080

Qy  361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln  380
Db 1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG  1140

Qy  381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyValLysLeuThr  400
Db 1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCAGGAGGTGAAACTGACC  1200

Qy  401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln  420
Db 1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCGTTGGCAGTTTTACGCAA  1260

Qy  421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHisHis  440
Db 1261 ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTGAGATCTCATCACCATCACCATCAC  1320

RESULT 8
ADO50301
ID   ADO50301 standard; DNA; 1323 BP.
XX
AC   ADO50301;
XX
DT   29-JUL-2004 (first entry)
XX
DE   Kangaroo rat Escherichia coli phytase 872PH1 mutant DNA.
XX
KW   Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW   phytate; animal feed; fish feed; dough; baking; mutant; gene; ds.
XX
OS   Escherichia coli.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..1323
FT                  /tag= a
FT                  /product= "Phytase mutant enzyme"
FT   unsure         214..216
FT                  /tag= b
```

```
FT          mutation    /note= "Encodes Arg"
FT                      replace(527,A)
FT                      /*tag= c
FT          mutation    replace(893,T)
FT                      /*tag= d
FT          mutation    replace(895,G)
FT                      /*tag= e
FT          mutation    replace(934,G)
FT                      /*tag= f
XX
PN   US2004091968-A1.
XX
PD   13-MAY-2004.
XX
PF   20-JUN-2003; 2003US-00601319.
XX
PR   13-AUG-1997; 97US-00910798.
PR   01-MAR-1999; 99US-00259214.
PR   13-APR-1999; 99US-00291931.
PR   25-MAY-1999; 99US-00318528.
PR   25-MAY-2000; 2000US-00580515.
PR   24-MAY-2001; 2001US-00866379.
XX
PA   (SHOR/) SHORT J M.
PA   (KRET/) KRETZ K.
PA   (GRAY/) GRAY K A.
PA   (BART/) BARTON N R.
PA   (GARR/) GARRETT J B.
PA   (ODON/) O'DONOGHUE E.
PA   (MATH/) MATHER E J.
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI   Mather EJ;
XX
DR   WPI; 2004-374952/35.
DR   P-PSDB; ADO50302.
XX
PT   Producing phytase, involves providing nucleic acid derived from bacteria
PT   encoding polypeptide having phytase activity, and expressing nucleic acid
PT   in yeast.
XX
PS   Disclosure; Page; 74pp; English.
XX
CC   The invention relates to a method for producing phytase that involves
CC   providing a nucleic acid encoding phytase derived from a bacteria, and
CC   expressing the nucleic acid in a yeast under conditions that allow
CC   expression of the enzyme in the yeast. The invention also relates to
CC   modified phytase enzyme which has improved thermal tolerance and protease
CC   stability at low pH. The phytase enzyme can be used in foodstuffs to
CC   improve the feeding value of phytate rich ingredients, and in diet of
CC   numerous animals including mammals, fowls and fishes, commercially
CC   significant mammals such as pigs, goats, laboratory rodents, commercially
CC   significant avian species such as chicken, ducks, doves, parrot, etc.,
CC   commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC   etc., in dough making and baking, in dietary aids for animals. The method
CC   provides easy manufacture of the active ingredient loaded biocompatible
CC   composition, higher yields and loading efficiency. The phytase
CC   incorporated in to the dietary aid is safe for animals. The present
CC   sequence is Kangaroo rat Escherichia coli phytase mutant DNA. Note: The
CC   present sequence is not shown in the specification but has been derived
CC   from Escherichia coli B phytase DNA ADO50291.
XX
SQ   Sequence 1323 BP; 325 A; 353 C; 356 G; 288 T; 0 U; 1 Other;
```

```
Qy    1  MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla   20
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  ATGAAAGCGATCTTAATCCCATTTTATCTCTTCGATTCCGTTAACCCGCAATCTGCA    60

Qy   21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly   40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT  120

Qy   41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA  180

Qy   61  ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu   80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGGTGAGCTGATTGCCTATCTC  240

Qy   81  GlyHisTyrArgGlnArgAlaArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro  100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  GGACATTACCACGCCAGCGTCTGGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG  300

Qy  101  GlnSerGlyGlnValAlaIleAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA  360

Qy  121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr  140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG  420

Qy  141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla  160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  TCCAGTCCGGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG  480

Qy  161  AsnValThrAspAlaIleLeuSerArgAlaAlaGlyGlySerIleAlaAspPheThrGlyHis  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  AACGTGACTGACGCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGGCTTTACCGGGCAT  540

Qy  181  ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys  200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC  600

Qy  201  LeuLysArgGluLysGlnAspGluSerLeuThrGlnAlaLeuProSerGluLeu  220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  CTTAAACGTGAGAAACAGGACGGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC  660

Qy  221  LysValSerAlaAspAsnAlaSerLeuThrGlyAlaAlaValSerLeuAlaSerMetLeuThr  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG  720

Qy  241  GluIlePheLeuLeuGlnIleGlnAlaGlnIleGlyMetProGluProGlyTrpGlyArgIleThr  260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  GAGATATTTTCTCCTGCAACAGCACAGGGAATGCCGGGGCCGGGGTGGGGAAGGATCACC  780

Qy  261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln  280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA  840

Qy  281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  CGCACGCCAGAGTTGGCCCGCAGCCGGCGCCACCCCGTTATTGATTTGATCAAGACAGCG  900

Qy  301  LeuThrProHisProProGlnLysGlnAlaThrGlyValThrLeuProThrSerValLeu  320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  TTGACGCCCCATCCACCGCAAAACAGGCGTATAGTGGACATTACCCACTTCAGTACTG  960

Qy  321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTCGAGCTCAACTGG 1020

Qy  341  ThrLeuProGlyGlnProAspAsnThrProProGlyGluLeuValPheGluArgTrp  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021  ACGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1080
```

Alignment Scores:

Pred. No.:	4.16e-200	Length:	1323
Score:	2279.00	Matches:	436
Percent Similarity:	99.1%	Conservative:	0
Best Local Similarity:	99.1%	Mismatches:	4
Query Match:	99.0%	Indels:	0
DB:		Gaps:	0

US-10-601-319-2 (1-440) x ADO50301 (1-1323)

```
Qy    361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CGTCGGCTAAGCCGATAAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140

Qy    381 GlnMetArgAspLysThrProLeuSerAsnThrProProGlyValValLysThrLeuThr 400
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCCGGAGAGGTGAAACTGACC 1200

Qy    401 LeuAlaGlyCysGluGluAsnAlaGlnAsnGlyMetCysSerLeuAlaGlyPheThrGln 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGCAGGATGTGAAGAGAATGCGCAGAGCGCATGTGTTCGTTGGCAGGTTTTACGCAA 1260

Qy    421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerLeuArgSerHisHisHisHisHis 440
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTGAGATCTCATCACCATCACCATCAC 1320

RESULT 9
ABK12514
ID    ABK12514 standard; DNA; 1299 BP.
XX
AC    ABK12514;
XX
DT    05-JUN-2002 (first entry)
XX
DE    DNA encoding phytase associated protein.
XX
KW    Phytase; gene; ds.
XX
OS    Unidentified.
XX
FH    Key    Location/Qualifiers
FT    CDS    1..1299
FT           /*tag= a
FT           /product= "Phytase associated protein"
XX
PN    KR99086028-A.
XX
PD    15-DEC-1999.
XX
PF    25-MAY-1998; 98KR-00018810.
XX
PR    25-MAY-1998; 98KR-00018810.
XX
PA    (WOOJ-) WOOJIN CO LTD.
XX
PI    Bae HD, Forceburgh CW, Goloben S, Cheng KJ;
XX
DR    WPI; 2000-645078/62.
DR    P-PSDB; AAU77775.
XX
PT    Novel phytase gene, recombinant phytase and usage thereof.
XX
PS    Claim 1; Fig 2; 10pp; Korean.
XX
CC    The invention relates to a novel phytase gene, a recombinant phytase gene
CC    and their uses. This sequence encodes a phytase associated protein,
CC    described in the invention
XX
SQ    Sequence 1299 BP; 317 A; 345 C; 355 G; 282 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1299
Score:        4.63e-196    Matches:           430
              2235.00      Conservative:        0
Percent Similarity:    99.5%    Mismatches:          0
Best Local Similarity:  99.5%    Indels:              2
Query Match:           97.1%    Gaps:                0
DB:                      3

US-10-601-319-2 (1-440) x ABK12514 (1-1299)

Qy      1 MetLysAlaIleLeuLeuProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTCTGATTCCCGTTAACCCCGCAATCTGCA 60
```

```
Qy   381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCGGGAGAGGTGAAACTGACC 1200

Qy   401 LeuAlaGlyCysGluGlyArgAsnAlaGlnAlaGlyMetCysSerLeuAlaGlyPheThrGln 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 CTGGCAGGATGTGAAGAGCAGAGCGCAGGGCATGTGTCGTTGCAGTTTTACCGCAA 1260

Qy   421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
         ||||||||||||||||||||||||||||||||||||
Db  1261 ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTG 1296
```

RESULT 10
ACA19297
ID ACA19297 standard; DNA; 1299 BP.
XX
AC ACA19297;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #954.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Escherichia coli.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2002; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L,   Zamudio C,   Malone C,   Haselbeck R,   Ohlsen KL,   Zyskind JW;
PI Wall D,   Trawick JD, Carr GJ,   Yamamoto R,   Forsyth RA,   Xu HH;
XX
DR WPI; 2003-029926/02.
DR P-PSDB; ABU15427.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 7167; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the extent
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
```

```
CC  strains; or (13) identifying the target of a compound that inhibits the
CC  proliferation of an organism. The antisense nucleic acids are useful for
CC  identifying proteins or screening for homologous nucleic acids required
CC  for cellular proliferation to isolate candidate molecules for rational
CC  drug discovery programs, or for screening homologous nucleic acids
CC  required for proliferation in cells other than S. aureus, S. typhimurium,
CC  K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC  prokaryotic essential genes. Note: The sequence data for this patent did
CC  not form part of the printed specification, but was obtained in
CC  electronic format directly from WIPO at
CC  ftp.wipo.int/pub/published_pct_sequences
XX
SQ  Sequence 1299 BP; 317 A; 345 C; 355 G; 282 T; 0 U; 0 Other;
```

Alignment Scores:
Pred. No.:
Score: 4.63e-196
Percent Similarity: 2235.00
Best Local Similarity: 99.5%
Query Match: 99.5%
DB: 97.1%
                           8

Length: 1299
Matches: 430
Conservative: 0
Mismatches: 2
Indels: 0
Gaps: 0

US-10-601-319-2 (1-440) x ACA19297 (1-1299)

```
Qy    1 MetLysAlaIleLeuIleLeuProPheLeuSerLeuLeuIleLeuProLeuThrProGlnSerAla 20
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCGATTCCGTTAACCCGCAATCTGCA 60

Qy   21 PheAlaGlnSerGlyGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGGTGATTGTCAGTCGTCATGT 120

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy   61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGCGGTGGTGAGCTAATCGCCTATCTC 240

Qy   81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysProPro 100
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG 300

Qy  101 GlnSerGlyGlnValAlaIleIleGluAlaAspValAspThrArgThrArgLysThrGlyGlu 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGTCTGGTCAGGTGCGGATTATTGCTGAGGCGTACCGTAAAACAGGCGGAA 360

Qy  121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GCCTTCGCGGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420

Qy  141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TCCAGTCCGGATCCGTTATTTAATCCTCTAAAAACTGGCGTTGTGCCAACTGGATAACCG 480

Qy  161 AsnValThrAspAlaAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 AACGTGACTGACCGGATCCTCAGCAGGCGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 540

Qy  181 ArgGlnThrAlaPheArgGluGluGlyValLeuAsnPheProGlnSerAsnLeuCys 200
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC 600

Qy  201 LeuLysArgGluLysGlnAspAlaAsplGlnSerCysSerSerLeuThrGlnAlaLeuProSerGluLeu 220
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CTTAAACGTGAGAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC 660

Qy  221 LysValSerAlaAspAsnValSerLeuThrThrGlyAlaValSerLeuAlaSerMetLeuThr 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 AAGGTGAGCGGCCGACAATGTCTCATTAACCGTGCGGTAAGCCTCGCCATCAATGCTGACG 720

Qy  241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGlyTrpProGlyArgIleThr 260
```

PI   Mrachko GT, Schellenberger V;
XX
DR   WPI; 2004-169498/16.
DR   P-PSDB; ADL16129.
XX
PT   Novel recombinant phytase having modified phytase activity comprising
PT   modification of amino acid residues in mature Escherichia coli phytase
PT   designated EBC18B2, useful as feed additive.
XX
PS   Claim 19; Fig 21; 107pp; English.
XX
CC   The invention relates to recombinant modified AppA phytases having
CC   modified phytase activity. The AppA phytases are derived from a mature
CC   Escherichia coli AppA phytase designated EBC18B2 (see ADL16120), and
CC   comprise subsitutions at least one of residues 26, 43, 46, 54, 73, 113,
CC   126, 184, 228, 384 or 410, or at a residue located within 5 residues of
CC   those listed above. The invention also encompasses polynucleotides
CC   encoding a modified AppA phytase and which further encodes a secretion
CC   signal sequence operable in Bacillus sp.; expression constructs, vectors
CC   and host cells (preferably Bacillus subtilise) comprising the
CC   polynucleotide sequences; and a method for the production of the modified
CC   AppA phytases. The invention also discloses a method of producing an
CC   enzyme having phytase activity; a method of producing a heterologous
CC   polypeptide having phytase activity in Bacillus subtilis; reducing the
CC   level of phosphorus in animal manure; a composition comprising a modified
CC   AppA phytase; and animal feed comprising cells, spores or plant parts,
CC   including seeds, capable of expressing a modified AppA phytase. The
CC   modified AppA phytases are useful as an animal feed additive,
CC   particularly in feeds for non-ruminant livestock such as pigs and poultry
CC   which lack the digestive enzymes for extracting phosphorus from phytate
CC   (myo-inositol hexakisphosphate), a substance found in plant seed which is
CC   an important source of phosphorus. The modified AppA phytases also have
CC   many other agricultural, industrial, medical and nutritional
CC   applications. For example, they can be used in the isolation and
CC   recovery of rare metals to produce lower phosphate homologs of phytate,
CC   which may be used in toothpaste and other dental care products as well as
CC   potential treatments or preventatives of bone resorption (e.g., in
CC   osteoporosis) and renal calculi (kidney stones); in food production
CC   (e.g., tofu and nutritional supplements); in wet grain milling; in
CC   cleaning, personal care, medical and nutritional products; in rust
CC   removal products; and in the removal of metals and other polycations from
CC   such diverse materials as waste products and carbonated beverages. The
CC   present sequence represents Escherichia coli strain K-12 AppA phytase
CC   gene.
XX
SQ   Sequence 1299 BP; 317 A; 345 C; 355 G; 282 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   Length: 1299
Score: 4.63e-196   Matches: 2235.00
Percent Similarity: 99.5%   Conservative: 0
Best Local Similarity: 99.5%   Mismatches: 2
Query Match: 97.1%   Indels: 0
DB: Gaps: 0

US-10-601-319-2 (1-440) x ADL16137 (1-1299)

Qy   1 MetLysAlaAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20

Db   1 ATGAAAGCGATCTTAATCCCATTTTATTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60

Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleSerArgHisGly 40

Db   61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60

Db   121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy   61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80

Db   181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGCGGTGGTGAGCTAATCGCCTATCTC 240

Db   721 GAGATATTTCTCCTGCAACAAGCACAGGAATGCCGGAGCCGGGGTGGGAGGATCACC 780

Qy   261 AspSerHisGlnTrpAsnThrLeuSerLeuLeuHisAsnAlaGlnPheTyrLeuLeuGln 280

Db   781 GATTCACACCAGTGGAACACCCTTGCTAAGTTTTGCATAACGCGCCAATTTTATTTGCTACAA 840

Qy   281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300

Db   841 CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCGTTATTAGATTTGATCAAGACAGCG 900

Qy   301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrGluProThrSerValLeu 320

Db   901 TTGACGCCCCATCCACCCAAAAACAGGCGTATGGTGACATTACCCACTTCAGTGCTG 960

Qy   321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340

Db   961 TTTATCGCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG 1020

Qy   341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluValPheGluValArgTrp 360

Db 1021 ACGCTTCCCGGTCAGCCGGATAACACGCCCCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1080

Qy   361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380

Db 1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140

Qy   381 GlnMetArgAspLysThrProLeuSerAsnThrProProGlyGluValLysLysLeuThr 400

Db 1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCCGGAGAGGTGAAACTGACC 1200

Qy   401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420

Db 1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA 1260

Qy   421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

Db 1261 ATCGTGAATGAAGCACGGATACCGGCGTGCAGTTTG 1296

RESULT 11
ADL16137
ID   ADL16137 standard; DNA; 1299 BP.
XX
AC   ADL16137;
XX
DT   06-MAY-2004 (first entry)
XX
DE   Escherichia coli K-12 AppA phytase gene.
XX
KW   AppA phytase; modified phytase; K-12; animal feed; feed additive;
KW   phosphorus; phytate; myo-inositol hexakisphosphate; food production;
KW   wet grain milling; dental care product; bone resorption; osteoporosis;
KW   osteopathic; kidney stone; metal removal; waste treatment;
KW   cleaning product; rust removal product; gene; ds.
XX
OS   Escherichia coli; K-12.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1299
FT   /*tag= a
FT   /product= "AppA phytase"
XX
PN   WO2004015084-A2.
XX
PD   19-FEB-2004.
XX
PF   11-AUG-2003; 2003WO-US025058.
XX
PR   12-AUG-2002; 2002US-0403330P.
XX
PA   (GEMV) GENENCOR INT INC.
XX
PI   Blattmann BO, Darzins A, Davis JM, Encell LP, Morrison TB;

```
AAD25463
ID   AAD25463 standard; DNA; 1901 BP.
XX
AC   AAD25463;
XX
DT   26-MAR-2002 (first entry)
XX
DE   Escherichia coli appA phytase wild type DNA.
XX
KW   Bacterial phytase; K12 appA phytase; protease stability; anabolic;
KW   gastrointestinal; nutritional value; feed treatment process; therapy;
KW   thermal tolerance; growth performance; alcoholic drink; biopulping;
KW   non-alcoholic drink; biobleaching; ds.
XX
OS   Escherichia coli.
XX
FH   Key             Location/Qualifiers
FT   CDS             188..1486
FT                   /*tag= a
FT                   /product= "E. coli appA phytase protein"
XX
PN   WO200190333-A2.
XX
PD   29-NOV-2001.
XX
PF   24-MAY-2001; 2001WO-US017118.
XX
PR   25-MAY-2000; 2000US-00580515.
XX
PA   (DIVE-) DIVERSA CORP.
XX
PI   Short JM,   Kretz KA,   Gray KA,   Barton NR,   Garrett JB,   O'donoghue E;
XX
DR   WPI; 2002-083108/11.
DR   P-PSDB; AAE15807.
XX
PT   New bacterial phytase for e.g. improving the nutritional value of phytate
PT   -containing foodstuffs and subsequently improving the growth performance
PT   of an organism that consumes it, or in treating animal digestive systems.
XX
PS   Claim 42; Fig 7; 170pp; English.
XX
CC   The patent discloses recombinant bacterial phytase from Escherichia coli
CC   K12 appA phytase. The enzyme has phytase activity and improved thermal
CC   tolerance when compared with wild-type phytase. It has improved protease
CC   stability at low pH. The recombinant phytase is useful for improving the
CC   nutritional value of phytate-containing foodstuffs and subsequently
CC   improving the growth performance of an organism that consumes it, in
CC   treating animal digestive systems, in feed treatment processes and for in
CC   vitro purposes related to research, discovery and development. They are
CC   also used for generating recombinant digestive system life forms, for
CC   producing or manufacturing alcoholic and non-alcoholic drinks based on
CC   the use of moulds, grains and/or plants, in biopulping and bio-bleaching
CC   where a reduction in the use of environmentally harmful chemicals that
CC   are traditionally used in the pulp and paper industry is desired and in
CC   the reduction or possible elimination of the need for mineral
CC   supplements, enzymes or therapeutic drugs for animals from the daily feed
CC   thus increasing the amount calories and nutrients present in the feed.
CC   The present sequence is a DNA encoding E. coli appA phytase wild type
CC   protein
XX
SQ   Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;
```

Alignment Scores:

```
Pred. No.:                               Length:         1901
Score:                7.83e-196          Matches:         430
Percent Similarity:   2235.00            Conservative:      0
Best Local Similarity: 99.5%             Mismatches:        2
Query Match:          99.5%              Indels:            0
DB:                   97.1%              Gaps:              0
                      6
```

US-10-601-319-2 (1-440) x AAD25463 (1-1901)

```
Qy       81 GlyHisTyrGlnArgGlyAsnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 GGACATTACCAACGCCAGCCGTCTGGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG 300

Qy      101 GlnSerGlyGlnValAlaIleIleAlaAlaAspValAspGluArgThrArgLysThrGlyGlu 120
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 CAGTCTGGTCAGGTCGCGGATTATTGCTGATGTCGACGAGCCGTACCGTAAAACAGGCGAA 360

Qy      121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 GCCTTCGCCGCCGGCTGGCACCTGACCTGTGCAATACCGTACCAGGCAGATACG 420

Qy      141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 TCCAGTCCGGATCCGTATTTAATCCTCTAAAAACTGGCGTTGCCAACTGGATAACGGCG 480

Qy      161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 AACGTGACTGACGCGATCCTCAGCAGGGCAGGGGTCAATTGCTGACTTTACCGGGCAT 540

Qy      181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 CGGCAAACGGCGTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600

Qy      201 LeuLysArgGlyGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 CTTAAACGTGGAGAAAAGCAGGAGCGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC 660

Qy      221 LysValSerAlaAlaAspAsnValSerLeuThrGlyAlaValAlaSerMetLeuThr 240
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 AAGGTGAGCGCGACAATGTCTCATTAACCGGTGCGTAAGGCCTGCATCAATGCTGACG 720

Qy      241 GluIlePheLeuLeuGlnGlnAsnAlaGlnIleGlyMetProGluProGlyTrpGlyArgIleThr 260
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCCGAGCGCGGGGTGGGGAAGGATCACC 780

Qy      261 AspSerHisGlnThrAsnThrLeuLeuSerLeuHisAsnAlaAlaGlnPheTyrLeuLeuGln 280
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCGCAATTTATTTGCTACAA 840

Qy      281 ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspPIleMetAlaAlaAla 300
                ||||||||| |||||||||||||||||||||||||| |||||||| ||||||||||||||||
Db      841 CGCACGGCCAGAGGTTGCCCGCCAGCCGCCGCACCCCGTTATTAGATTTGATCAAGACAGCG 900

Qy      301 LeuThrProHisProProGlnLysProGlnAlaAlaTyrArgGlyValThrProThrSerValLeu 320
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 TTGACGCCCCATCCACCCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCTG 960

Qy      321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 TTTATCGGCCGGACACGATACTAATCTGGCAAATCTCGGAGCGCACTGGAGCTCAACTGG 1020

Qy      341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyValPheGluGluArgTrp 360
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 ACGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1080

Qy      361 ArgArgLeuSerAspAsnSerGlnAsnTrpIleSerValSerLeuValPheGlnThrLeuGln 380
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 CGTCGGCTAAGCGATAACAGCCAGTGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140

Qy      381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyProGlyValValLeuThr 400
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGGCCGCCGGAGAGGGTGAAACTGACC 1200

Qy      401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA 1260

Qy      421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
                ||||||||||||||||||||||||||||||||
Db     1261 ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTG 1296
```

RESULT 12

```
Qy    1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla   20
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  188 ATGAAAGCCGATCTTAATCCCATTTTAATCTCTTCTGATTCCGTTAACCCGCAATCTGCA  247

Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly   40
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  248 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGTCGAAAGTGTCGAATTGTCAGTCGTCATGGT  307

Qy   41 ValArgAlaProThrLysMetGlnAlaThrArgGlyAspValThrProAlaTrpPro   60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  308 GTGCGTGCTCCAACCAAGGCCACGAGCAATGCGAGGATGTCACCCCAGACGCATGGCCA  367

Qy   61 ThrTrpProValLysGlyLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu   80
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  368 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCCGCNGGTGGTGAGCTAATCGCCTATCTC  427

Qy   81 GlyHisTyrGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro  100
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  428 GGACATTACCAACGCCAGCCGTCTCGGTAGCCGACCGGATTCGTCGGCGAAAAAGGGCTGCCCG  487

Qy  101 GlnSerGlyGlnValAlaIleIleAlaAspValArgThrArgLysThrGlyGlu  120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  488 CAGTCTGGTCAGGTGCGATTATTGCTGATGTCGACAGAGCGTACCCGTAAAACAGGCGAA  547

Qy  121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr  140
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  548 GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG  607

Qy  141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla  160
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  608 TCCAGTCCGGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG  667

Qy  161 AsnValThrAspAlaIleAlaLeuGluAlaArgAlaGlyGlySerIleAlaAspAspPheHisGlyHis  180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  668 AACGTGACTGACGCGATCGCTCAGGAGCGCAGGAGGGCAGGGGTCAATTGCTGACTTTACCGGGCAT  727

Qy  181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys  200
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  728 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC  787

Qy  201 LeuLysArgGlnLysGlnAspAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  788 CTTAAACGTGAGAAACAGGACGACGAAAGCTGTTCATTAACGCAGGCATTACCATCCGGAACTC  847

Qy  221 LysValSerAlaAspAspAsnAlaValSerLeuThrArgGlyAlaAlaValAlaSerMetLeuThr  240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  848 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG  907

Qy  241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyTrpGlyArgIleThr  260
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  908 GAGATATTTCTCTGCAACAAGCACAGGGAATGCCGGAGCCCGGCGGTGGGGAAGGATCACC  967

Qy  261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln  280
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  968 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATACGCGCAATTTTATTTGCTACAA  1027

Qy  281 ArgThrProGluGluValValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla  300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1028 CGCACGCCAGAGGTTGCCCGCAGCCCGCGCCACCCGTTATTAGATTTGATCAAGACAGCG  1087

Qy  301 LeuThrProHisProProGlnLysGlnIleAlaTyrGlyValThrLeuProThrSerValLeu  320
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1088 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCTG  1147

Qy  321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1148 TTTATCGCCGGACACGATACTAATCTCGGCAAATCTCGGCGGCCCACTGGAGCTCAACTGG  1207

Qy  341 ThrLeuProGlyTyrGlnProProAspThrProProGlyTyrGluLeuValPheGluLeuArgTrp  360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1208 ACGCTTCCCGGTCAGCCGGATAACACCGCCGCCAGGTGCGTGAACTGGTGTTTGAACGGCTGG  1267

Qy  361 ArgArgLeuSerAspAsnSerGlnAsnTrpIleGlnIleValSerLeuLeuValThrLeuGln  380
```

```
Db 1268 CGTCGGCTAAGCAGCATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG  1327

Qy  381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProGluValLysLeuThr   400
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1328 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGGAGGTGAAACTGACC  1387

Qy  401 LeuAlaGlyCysGluGlnArgAsnAlaGlnMetCysGlySerLeuAlaGlyPheThrGln  420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1388 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTGTGTCGTTGGCAGGTTTTACGCAA  1447

Qy  421 IleValAsnGluValAlaArgIleProAlaCysSerLeu  432
        |||||||||||||||||||||||||||||||||||||||
Db 1448 ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTG  1483
```

RESULT 13
ADA19449
ID   ADA19449 standard; DNA; 1901 BP.
XX
AC   ADA19449;
XX
DT   20-NOV-2003   (first entry)
XX
DE   E. coli K12 appA gene encoding phytase.
XX
KW   Phytase; ds; gene; phytate; appA gene; animal feed; inorganic phosphate;
KW   digestion enhancement; transgenic; thermal tolerance; protease stability.
XX
OS   Escherichia coli; strain K12.
XX
FH   Key   Location/Qualifiers
FT   CDS   188..1486
FT   /*tag= a
FT   /product= "Phytase"
XX
PN   US2002136754-A1.
XX
PD   26-SEP-2002.
XX
PF   24-MAY-2001; 2001US-00866379.
XX
PR   13-AUG-1997;   97US-00910798.
PR   01-MAR-1999;   99US-00259214.
PR   13-APR-1999;   99US-00291931.
PR   25-MAY-1999;   99US-00318528.
PR   25-MAY-2000;   2000US-00580515.
XX
PA   (SHOR/)   SHORT J M.
PA   (KRET/)   KRETZ K A.
PA   (GRAY/)   GRAY K A.
PA   (BART/)   BARTON N R.
PA   (GARR/)   GARRETT J B.
PA   (DONO/)   O' DONOGHUE E.
PA   (MATH/)   MATHUR E J.
XX
PI   Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB;
PI   O' Donoghue E, Mathur EJ;
XX
DR   WPI; 2003-040002/03.
DR   P-PSDB; ADA19450.
XX
PT   Isolated Escherichia coli polynucleotide encoding a modified phytase
PT   enzyme, useful in the production of animal feed, for improving the
PT   nutritional value of phytate-containing foodstuff and for enhancing
PT   digestion in humans and animals.
XX
PS   Claim 42; Fig 7; 62pp; English.
XX
CC   The invention relates to an isolated Escherichia coli polynucleotide
CC   encoding a phytase enzyme appearing as ADA19450 and having amino acids
CC   modification from W68B, Q84W, A95P, K97C, S168E, R180Y, N226C or Y277D.
CC   Also included the E. coli appA gene ADA19449 (or an oligonucleotide
CC   derived from it) or its mutant sequence ADA19452, expression vectors,

```
CC   host cells, a method of improving nutritional value of a phytate-
CC   containing foodstuff by contacting the phytate-containing foodstuff with
CC   a pure phytase enzyme mutated phytase (where the phytase enzyme catalyzes
CC   the liberation of inorganic phosphate from the phytate in the phytate-
CC   containing foodstuff), a method to produce an animal feed containing a
CC   microbial phytase (comprising culturing the plant cell, plant part or
CC   plant under conditions where the nucleotide sequence is expressed and
CC   converting the plant cells, plants or plant into a composition for animal
CC   feed), a feed composition for animals (comprising the plant seeds, plant
CC   cells, plant parts or plants in admixture with a phytate-containing
CC   foodstuff), a method to treat a human or an animal able to benefit from
CC   digestive enhancement by the activity of an exogenous phytase enzymes
CC   comprising administering to the human or animal the plant seed, plant
CC   cells, plant parts or plants of a transgenic plant which is modified to
CC   contain an expression system which expresses a nucleotide sequence
CC   encoding a phytase enzyme, a transgenic non-human organism whose genome
CC   comprising a heterologous nucleic acid sequence encoding a polypeptide
CC   having phytase activity. The phytase enzyme is useful for improving the
CC   nutritional value of phytate-containing foodstuff, in the production of
CC   animal feed and for enhancing digestion in humans and animals. The
CC   invented method improves thermal tolerance and protease stability. It
CC   also improves the feeding value of phytate rich ingredients. The present
CC   sequence represents the E. coli K12 appA gene encoding wild-type phytase.
XX
SQ   Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:              7.83e-196        Length:           1901
Score:                  2235.00          Matches:          430
Percent Similarity:     99.5%            Conservative:     0
Best Local Similarity:  99.5%            Mismatches:       2
Query Match:            97.1%            Indels:           0
DB:                     8                Gaps:             0

US-10-601-319-2 (1-440) x ADA19449 (1-1901)

Qy    1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  188 ATGAAAGCGATCTTAATCCCATTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 247

Qy   21 PheAlaGlnSerGluProGluLeuLysGluSerValValIleSerValSerArgHisGly 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  248 TTCGCTCAGAGTGAGCCGGAGCTGAAGGTGAAAGTGTGGTGATTGTCAGTCGTCATGGT 307

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 367

Qy   61 ThrTrpProValLysGluGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  368 ACCTGGCCGGTAAAACCGGTGGCTGACACACCGGGNGGTGTGGTGAGCTAATCGCCTATCTC 427

Qy   81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  428 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGCTGGCGAAAAAGGGGCTGCCCG 487

Qy  101 GlnSerGlyGlnValAlaIleIleAlaIleIleAlaAlaAspValAspArgThrArgLysThrGlyGlu 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  488 CAGTCTGGTCAGGTCGCGATTATTGCTGCGATTGTGCGACGAGCGTACCCGTAAAACAGGCGAA 547

Qy  121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  548 GCCTTCGCCGCGGGCCTGGCCACCTGACGTGCAATAACCGTCCAATACCCCAGGCAGATACG 607

Qy  141 SerSerProAspPheAsnProLeuLysThrGlyValCysGlnLeuAspAspAsnAla 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  608 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGGCGGCGGTTGCCAACTGGATAACGCG 667

Qy  161 AsnValThrAspAlaAlaIleLeuSerArgAlaGlyGlySerIleAlaAlaAspPheThrGlyHis 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  668 AACGTGACTGACGCGGCCATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTTACCGGGCAT 727

Qy  181 ArgGlnThrAlaPheArgGluGluLeuGlyValLeuAsnPheProGlnSerAsnLeuCys 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  728 CGGCAAACGGCGGTTTCGGGAACTGGAAGCGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 787

Qy  201 LeuLysArgLysGlnAspGluSerLeuThrGlnAlaLeuProSerGluLeu 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  788 CTTAAACGGTGAGAAACAGGACGAAAGTCTTCATTAACGCCAGGCATTACCATCGGACTC 847

Qy  221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  848 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCCGCATCAATGTGACG 907

Qy  241 GluIlePheLeuLeuGlnIleAlaGlnGlyMetProGluProGlyTyrProGlyArgIleThr 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  908 GAGATATTTCCTGCACAAGCACAGGAATGCCGGAGCCGGGTGGGAAGGATCACC 967

Qy  261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  968 GATTCACCACCAGTGGACACCCTTGCTAAGTTTGCATAACGGGCCAATTTATTTGCTACAA 1027

Qy  281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1028 CGCACGCCAGAGGTTGCCCGGCCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 1087

Qy  301 LeuThrProHisProProGlnLysGlnAlaThrLeuAlaTyrGlyValThrLeuProThrSerValLeu 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1088 TTGACGCCCCATCCACCGCAAAAACAGGCGATAGTGTGACATTACCCACTTCAGTGCTG 1147

Qy  321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1148 TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGCACTGGAGCTCAACTGG 1207

Qy  341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1208 ACGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1267

Qy  361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1268 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1327

Qy  381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1328 CAGATGCGTGATAAAACGCCCGTGTCATTAAATACCCCGCCGGGAGAGGTGAAACTGACC 1387

Qy  401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1388 CTGGCAGGATGTGAAGAGCGAAATGCCAGGGCATGTGTTCGTTGGCAGGTTTTTACGCAA 1447

Qy  421 IleValAlaAsnGluAlaArgIleProAlaCysSerLeu 432
          ||||||||||||||||||||||||||||||||||||||||
Db 1448 ATCGTGGATGAATGAAGCACGCATACCGGCGTGCAGTTTG 1483

RESULT 14
ADC87744
ID   ADC87744 standard; DNA; 1901 BP.
XX
AC   ADC87744;
XX
DT   01-JAN-2004 (first entry)
XX
DE   DNA encoding Escherichia coli appA phytase.
XX
KW   Phytase; food supplement; enzyme delivery matrix; soybean meal;
KW   thermotolerance; thermostability; kernel; phytate;
KW   myo-inositol-hexaphosphate; inositol; inorganic phosphate;
KW   thermotolerant; feed value; digestion; gene; ds; appA.
XX
OS   Escherichia coli.
XX
FH   Key             Location/Qualifiers
FT   CDS             188..1486
FT                   /tag= a
FT                   /product= "AppA phytase"
FT                   /transl_except= (pos:401..403,aa:Arg)
```

XX
PN   US2003103958-A1.
XX
PD   05-JUN-2003.
XX
PP   24-MAY-2002; 2002US-00156660.
XX
PR   13-AUG-1997;   97US-00910798.
PR   01-MAR-1999;   99US-00259214.
PR   13-APR-1999;   99US-00291931.
PR   25-MAY-1999;   99US-00318528.
PR   25-MAY-2000;   2000US-00580515.
PR   24-MAY-2001;   2001US-00866379.
XX
PA   (DIVE-) DIVERSA CORP.
XX
PI   Short JM,   Kretz K,   Gray KA,   Barton NR,   Garrett JB,   O'donoghue B;
PI   Mathur EJ;
XX
DR   WPI, 2003-787039/74.
DR   P-PSDB; ADC87745.
XX
PT   New nucleic acid encoding a polypeptide having phytase activity, useful
PT   in improving the feeding value of phytate rich ingredients or as an aid
PT   in phytate digestion.
XX
PS   Example 1; SEQ ID NO 3; 113pp; English.
XX
CC   The invention discloses a new isolated or recombinant nucleic acid which
CC   encodes a polypeptide having a phytase activity. Also claimed is a
CC   nucleic acid probe, an amplification primer sequence pair, an expression
CC   cassette comprising the nucleic acid, a vector comprising the nucleic
CC   acid, a transgenic non-human animal or plant, or its seed, comprising the
CC   nucleic acid, an antisense oligonucleotide, inhibiting the translation of
CC   a phytase message in a cell, a heterodimer comprising the polypeptide and
CC   a second domain, an array comprising immobilised polypeptide or nucleic
CC   acid, a hybridoma comprising an antibody that specifically binds to the
CC   polypeptide, a food supplement for an animal, an edible enzyme delivery
CC   matrix, an edible pellet comprising a granule edible carrier and the
CC   polypeptide, a feed composition, a soybean meal, isolating or identifying
CC   the polypeptide, making an anti-phytase antibody, producing a recombinant
CC   polypeptide, determining whether a compound binds to the polypeptide,
CC   identifying a modulator, whole cell engineering of new or modified
CC   phenotypes by using real-time metabolic flux analysis, increasing
CC   thermotolerance or thermostability of the phytase polypeptide, increasing
CC   the resistance of the phytase polypeptide to enzymatic inactivation in a
CC   digestive system and processing of corn and sorghum kernels. The phytase
CC   activity comprises catalysis of phytate (myo-inositol-hexaphosphate) to
CC   inositol and inorganic phosphate and is thermotolerant. The nucleic acid
CC   is useful in improving the feeding value of phytate rich ingredients or
CC   as an aid in phytate digestion. The sequence presented is the DNA
CC   encoding the Escherichia coli appA phytase.
XX
SQ   Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:
Score: 7.83e-196
Percent Similarity: 2235.00
Best Local Similarity: 99.5%
Query Match: 99.5%
DB: 97.1%
                      10

Length: 1901
Matches: 430
Conservative: 0
Mismatches: 2
Indels: 0
Gaps: 0

US-10-601-319-2 (1-440) x ADC87744 (1-1901)

```
Qy    1  MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla  20
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  188  ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCTTAACCCCGCAATCTGCA  247

Qy   21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly  40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  248  TTCGCTCAGAGTGAGCCGAGCTGAAGCTGGAAAGTGTGGTTGATTGTCAGTCGTCATGGT  307
```

```
Qy   41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro  60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  308  GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA  367

Qy   61  ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGluLeuIleAlaTyrLeu  80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  368  ACCTGGCCCGTAAAACTGGGTTGGCTGACACCGCGNGGTGTGAGCTAATCGCCTATCTC  427

Qy   81  GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro  100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  428  GGACATTACCAACGCCAGGGTCTCGTGTAGCCGACGACCGGATTGCTGGGCGAAAAAGGGCTGCCCG  487

Qy  101  GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  488  CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA  547

Qy  121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr  140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  548  GCCTTCGCCGCCGGGCTGGCACCTGACTGCGCAATAACCGTACATACCCAGGCAGATACG  607

Qy  141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla  160
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  608  TCCAGTCCGGATCCGTTATTTAATCCTCTAAAAACTGGGGTTTGCCAACTGGATAACGCG  667

Qy  161  AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  668  AACGTGACTGACGCGATCCTCAGCAGGCGAGGGGTCAATTGCTGACTTTACCGGGCAT  727

Qy  181  ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnProGlnSerAsnLeuCys  200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  728  CGGCAAACGGCGTTCGCGAACTGGAACGGGTGCTTAATTTCCGCAATCAAACTGTGC  787

Qy  201  LeuLysArgGluLysLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  788  CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGGCCATTACCATCGGAACTC  847

Qy  221  LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  848  AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCCATCAATGCTGACG  907

Qy  241  GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr  260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  908  GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGTGGGGAAGGATCACC  967

Qy  261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln  280
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  968  GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA  1027

Qy  281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1028  CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG  1087

Qy  301  LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu  320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1088  TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCTG  1147

Qy  321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1148  TTTATCGCCGGACACGATACTAATCTCGGCAAATCTCGGAGTCCAACTGG  1207

Qy  341  ThrLeuProGlyGlnProLeuAsnThrProProGlyGlyGluLeuValPheGluArgTrp  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1208  ACGCTTCCCGGTCAGCCGCTGAATACGCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG  1267

Qy  361  ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln  380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1268  CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG  1327

Qy  381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr  400
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1328  CAGATGCGTGATAAAACGCCGCTGTCATTAAATACCGCCCCGGAGAGGTGAAACTGACC  1387

Qy  401  LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln  420
```

```
Db   1388 CTGGCAGGATGTGAAGAGCGAAATGCCCAGGCATGTGTTCGTTCGGCCAGGTTTTACGCAA 1447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    421 IleValAlaArgIleProAlaCysSerLeu 432
          ||||||||||||||||||||||||
Db   1448 ATCGTGAATGAAGCACGCATACCGGGCTGCAGTTTG 1483


RESULT 15
ADO50297
ID   ADO50297 standard; DNA; 1901 BP.
XX
AC   ADO50297;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Escherichia coli K12 appA phytase DNA.
XX
KW   AppA phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW   phytate; animal feed; fish feed; dough; baking; gene; ds.
XX
OS   Escherichia coli; K12.
XX
FH   Key            Location/Qualifiers
FT   CDS            188..1486
FT                  /*tag= a
FT                  /product= "AppA phytase"
FT   unsure         401..403
FT                  /*tag= b
FT                  /note= "Encodes Arg"
XX
PN   US200409196B-A1.
XX
PD   13-MAY-2004.
XX
PF   20-JUN-2003; 2003US-00601319.
XX
PR   13-AUG-1997;  97US-00910798.
PR   01-MAR-1999;  99US-00259214.
PR   13-APR-1999;  99US-00291931.
PR   25-MAY-1999;  99US-00318528.
PR   25-MAY-2000;  2000US-00580515.
PR   24-MAY-2001;  2001US-00863379.
XX
PA   (SHOR/) SHORT J M.
PA   (KRET/) KRETZ K.
PA   (GRAY/) GRAY K A.
PA   (BART/) BARTON N R.
PA   (GARR/) GARRETT J B.
PA   (ODON/) O'DONOGHUE E.
PA   (MATH/) MATHER E J.
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI   Mather EJ;
XX
DR   WPI; 2004-374952/35.
DR   P-PSDB; ADO50298.
XX
PT   Producing phytase, involves providing nucleic acid derived from bacteria
PT   encoding polypeptide having phytase activity, and expressing nucleic acid
PT   in yeast.
XX
PS   Disclosure; SEQ ID NO 7; 74pp; English.
XX
CC   The invention relates to a method for producing phytase that involves
CC   providing a nucleic acid encoding phytase derived from a bacteria, and
CC   expressing the nucleic acid in a yeast under conditions that allow
CC   expression of the enzyme in the yeast. The invention also relates to
CC   modified phytase enzyme which has improved thermal tolerance and protease
CC   stability at low pH. The phytase enzyme can be used in foodstuffs to
CC   improve the feeding value of phytate rich ingredients, and in diet of
CC   numerous animals including mammals, fowls and fishes, commercially
CC   significant mammals such as pigs, goats, laboratory rodents, commercially
CC   significant avian species such as chicken, ducks, doves, parrot, etc.,
```

```
CC   commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC   etc., in dough making and baking, in dietary aids for animals. The method
CC   provides easy manufacture of the active ingredient loaded biocompatible
CC   composition, higher yields and loading efficiency. The phytase
CC   incorporated in to the dietary aid is safe for animals. The present
CC   sequence is Escherichia coli K12 appA phytase DNA.
XX
SQ   Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;
```

```
Alignment Scores:
Pred. No.:                                        1901
Score:                      7.83e-196             430
Percent Similarity:         2235.00
Best Local Similarity:      99.5%            Matches:       430
Query Match:                99.5%            Conservative:    0
DB:                         97.1%            Mismatches:      2
                                             Indels:          0
                                             Gaps:            0
```

```
US-10-601-319-2 (1-440) x ADO50297 (1-1901)

Qy     1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 ATGAAAGCGATCTTAATCCCATTTTATCTCTTCTGATTCCTTTAACCCCGCAATCTGCA 247

Qy    21 PheAlaGlnSerGluProGluLysGluSerValValIleValSerArgHisGly 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGGTGATCGTCAGTCGTCATGGT 307

Qy    41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 367

Qy    61 ThrTrpProValLysGlyGlyTrpLysThrProArgGlyGlyGluIleLeuIleAlaTyrLeu 80
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNNGGTGGTGAGCTAATGCCTATCTC 427

Qy    81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyGlyCysPro 100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGCTGGCGAAAAGGGCTGCCCG 487

Qy   101 GlnSerGlyGlnValAlaIleIleAlaAspGlyLeuLeuAlaLysGlyGlyGlu 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488 CAGTCTGGTCAGGTCGCGATTATTGCTGATGGTCGACGAGCGTACCGTAAAACAGGCGAA 547

Qy   121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548 GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 607

Qy   141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   608 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG 667

Qy   161 AsnValThrArgAspAlaIleLeuSerArgAlaAlaGlyGlySerIleAlaAspPheThrGlyHis 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668 AACGTGACTAGGGATGCGATCCTCAGCAGCGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 727

Qy   181 ArgGlnThrAlaPheArgGluArgLeuValLeuAsnPheProGlnSerAsnLeuCys 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   728 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 787

Qy   201 LeuLysArgGluLysGlyAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   788 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCAATTACCATCCGGAACTC 847

Qy   221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerAlaSerMetLeuThr 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   848 AAGGTGAGCGCCGACAATGTCATTAACCGGTGACGTAAGCCTGCATCAATGCTGACG 907

Qy   241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrProGlyArgIleThr 260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   908 GAGATATTTCTCTGCAACAAGCACAGGAATGCCCGGAGCCCGGGCTGGGAAGGGATCACC 967

Qy   261 AspSerHisGlnTrpAsnThrLeuLeuGlnHisAsnAlaGlnPheTyrLeuLeuGln 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    968  GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCCAATTTTATTTGCTACAA  1027

Qy    281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla   300
                ||||||||||||||||||||||||||||||||||||||||||||||      |||
Db   1028  CGCACGCCAGAGGTTGCCCGCCAGCCGCCACCCCGTTATTAGATTTGATCAAGACAGCG   1087

Qy    301  LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu   320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088  TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCTG   1147

Qy    321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp   340
                |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148  TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCCACTGGAGCTCAACTGG   1207

Qy    341  ThrLeuProGlyGlnProAspAsnThrProProGlyGlyLeuValValPheGluGluArgTrp   360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208  ACGCTTCCCGGTCAGCCGGATAACACGCCCGGAGGGTGGTGAACTGGTNGTTTGAACGGCTGG   1267

Qy    361  ArgArgLeuSerAspAsnSerGlnIleGlnIleValSerLeuValPheGlnThrLeuGln   380
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268  CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG   1327

Qy    381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValValLysLeuThr   400
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328  CAGATGCGTGATAAAACCCGCTGTCATTAAATACGCCGCCCGAGAGGTGAAACTGACC   1387

Qy    401  LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln   420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388  CTGGCAGGATGTGAAGAGCGAAATGCCGCAGGGCATGTGTTCGTTGGCAGGTTTACGCCAA   1447

Qy    421  IleValAsnGluAlaArgIleProAlaCysSerLeu   432
                ||||||||||||||||||||||||||||||||||
Db   1448  ATCGTGAATGAAGCACGCATACCGGCCGTGCAGTTTG   1483
```

Search completed: June 14, 2006, 10:57:59
Job time : 591.615 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.9

Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 14, 2006, 10:58:44 ; Search time 5255.78 Seconds
           (without alignments)
           8030.264 Million cell updates/sec

Title: US-10-601-319-2
Perfect score: 2302
Sequence: 1 MKAILIPPLSLLIPLTPQSA.........IVNEARIPACSLRSHHHHHH 440

Scoring table: BLOSUM62
           Xgapop 10.0 , Xgapext 0.5
           Ygapop 10.0 , Ygapext 0.5
           Fgapop   6.0 , Fgapext 7.0
           Delop   6.0 , Delext   7.0

Searched: 6366136 seqs, 3197710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10601319/runat_12062006_180050_27480/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss06h
-USER=US10601319 @CGN_1_1.3991 @runat_12062006_180050_27480 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :   GenEmbl:*
            1: gb_env:*
            2: gb_pat:*
            3: gb_ph:*
            4: gb_pl:*
            5: gb_pr:*
            6: gb_ro:*
            7: gb_sts:*
            8: gb_sy:*
            9: gb_un:*
           10: gb_vi:*
           11: gb_ov:*
           12: gb_htg:*
           13: gb_in:*
           14: gb_om:*
           15: gb_ba:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2302	100.0	1323	2	AR108133	AR108133 Sequence
2	2302	100.0	1323	2	AR127818	AR127818 Sequence
3	2302	100.0	1323	2	AR130956	AR130956 Sequence
4	2302	100.0	1323	2	AR492966	AR492966 Sequence
5	2302	100.0	1323	2	AR636185	AR636185 Sequence
6	2302	100.0	1323	2	AX052713	AX052713 Sequence
7	2302	100.0	1323	2	AX338966	AX338966 Sequence
8	2302	100.0	1323	2	AX356566	AX356566 Sequence
9	2235	97.1	1901	2	AR636190	AR636190 Sequence
10	2235	97.1	1901	2	AX356572	AX356572 Sequence
11	2235	97.1	3470	15	ECOAPPAA	M58708 Escherichia
12	2235	97.1	3470	2	AX042376	AX042376 Sequence
13	2235	97.1	4060	2	AX042374	AX042374 Sequence
14	2235	97.1	6116	2	AX042375	AX042375 Sequence
15	2235	97.1	6708	2	AX042373	AX042373 Sequence
16	2235	96.9	17732	2	AX042378	AX042378 Sequence
17	2235	97.1	20623	2	AX042372	AX042372 Sequence
18	2235	97.1	110000	15	U00096_10	Continuation (11 o
19	2235	97.1	110000	15	AP009048_10	Continuation (11 o
20	2235	96.9	110000	15	CP000038_10	Continuation (11 o
21	2231	96.9	1315	15	AF537219	AF537219 Escherich
22	2229	96.8	1901	15	L03370	L03370 Escherichia
23	2229	96.8	1901	15	ECOACPHC	L03372 Escherichia
24	2229	96.8	1901	15	ECOACPHD	L03373 Escherichia
25	2229	96.8	1901	15	ECOACPHE	L03374 Escherichia
26	2228	96.8	1901	15	ECOACPHB	L03371 Escherichia
c 27	2228	96.8	110000	15	CP000036_22	Continuation (23 o
28	2226	96.7	1901	2	AR636191	AR636191 Sequence
29	2226	96.7	1901	2	AX356574	AX356574 Sequence
30	2225	96.7	1901	15	ECOACPHF	L03375 Escherichia
31	2222	96.5	110000	15	AE005674_10	Continuation (11 o
32	2221	96.5	1489	15	AR277753	AR277753 Escherich
33	2221	96.5	1489	15	AY496073	AY496073 Escherich
34	2221	96.5	1901	2	AR636189	AR636189 Sequence
35	2221	96.5	1901	2	AX356571	AX356571 Sequence
36	2220	96.4	1901	2	AR636188	AR636188 Sequence
37	2220	96.4	1901	2	AX356570	AX356570 Sequence
38	2218	96.4	110000	15	AE014073_10	Continuation (11 o
39	2217	96.3	1486	2	AR630587	AR630587 Sequence
40	2213	96.1	110000	15	AE014075_10	Continuation (11 o
41	2210	96.0	110000	15	AE005174_12	Continuation (13 o
42	2210	96.0	110000	15	AE005174_13	Continuation (14 o
43	2210	96.0	110000	15	BA000007_12	Continuation (13 o
44	2205	95.8	5421	2	AX042377	AX042377 Sequence
45	2197	95.4	1486	2	AR630588	AR630588 Sequence

RESULT 1
AR108133
LOCUS   AR108133   1323 bp   DNA   linear   PAT 14-FEB-2001
DEFINITION Sequence 1 from patent US 6110719.
ACCESSION   AR108133
VERSION   AR108133.1 GI:12823620
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1323)
  AUTHORS   Kretz,K.
  TITLE   Phytase
  JOURNAL   Patent: US 6110719-A 1 29-AUG-2000;
FEATURES   Location/Qualifiers
     source   1..1323
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

ALIGNMENTS

Alignment Scores:
Pred. No.:   8.92e-205   Length: 1323
Score: 2302.00   Matches: 440
Percent Similarity: 100.0%   Conservative: 0
Best Local Similarity: 100.0%   Mismatches: 0
Query Match: 100.0%   Indels: 0
DB: 2   Gaps: 0

```
                                                              1021 ACGCTTCCGGTCAGCCGGATAACACGCGCCAGGTGGTGACTGGTGTTGAACGCTGG 1080
Db

Qy   361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnIleValSerLeuValPheGlnThrLeuGln 380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTCAGTTTCCGTGTCTTCCAGACTTTACAG 1140

Qy   381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyValValLysLeuThr 400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCCGGAGAGGTGAAACTGACC 1200

Qy   401 LeuAlaGlyCysGluGluArgAsnAlaGlnAlaGlyMetCysSerLeuAlaGlyPheThrGln 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGCGGGCATGTGTTCGTTGGCAGGTTTTACGCAA 1260

Qy   421 IleValAsnGluAlaArgIleProAlaCysSerLeuAlaArgSerHisHisHisHisHis 440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320
```

```
RESULT 2
AR127818
LOCUS          AR127818                1323 bp    DNA        linear     PAT 16-MAY-2001
DEFINITION     Sequence 1 from patent US 6183740.
ACCESSION      AR127818
VERSION        AR127818.1  GI:14115480
KEYWORDS       .
SOURCE         Unknown.
  ORGANISM     Unknown.
               Unclassified.
REFERENCE      1  (bases 1 to 1323)
  AUTHORS      Short,J.M. and Kretz,K.A.
  TITLE        Recombinant bacterial phytases and uses thereof
  JOURNAL      Patent: US 6183740-A 1 06-FEB-2001;
FEATURES             Location/Qualifiers
     source          1..1323
                     /organism="unknown"
                     /mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.:                          Length:           1323
Score:               8.92e-205      Matches:          440
Percent Similarity:  2302.00        Conservative:     0
Best Local Similarity: 100.0%       Mismatches:       0
Query Match:         100.0%         Indels:           0
DB:                  2              Gaps:             0
```

```
US-10-601-319-2 (1-440) x AR127818 (1-1323)

Qy     1 MetLysAlaIleLeuIleLeuProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAAAGCGATCTTAATCCTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60

Qy    21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerValArgHisGly 40
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGGTGATTGTCAGTGTCGTCATGGT 120

Qy    41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValAspThrAlaTrpPro 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy    61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC 240

Qy    81 GlyHisTyrGlnAsnArgLeuValAlaAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGACATTACCAACGCCAGCGGTCGGTAGCGCAGATTTGGTGCGAAAAAGGGCTGCCCG 300

Qy   101 GlnSerGlyGlnValAlaIleIleAlaIleIleAlaAspValAspArgThrArgLysThrGlyGlu 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGTCTGGTCAGGTCGCGGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA 360

Qy   121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
US-10-601-319-2 (1-440) x AR108133 (1-1323)

Qy     1 MetLysAlaIleLeuIleLeuProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAAAGCGATCTTAATCCTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60

Qy    21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerValArgHisGly 40
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGGTGATTGTCAGTGTCGTCATGGT 120

Qy    41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValAspThrAlaTrpPro 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy    61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC 240

Qy    81 GlyHisTyrGlnAsnArgLeuValAlaAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGACATTACCAACGCCAGCGGTCGGTAGCGCAGATTGCTGGGTGCGAAAAAGGGCTGCCCG 300

Qy   101 GlnSerGlyGlnValAlaIleIleAlaIleIleAlaAspValAspArgThrArgLysThrGlu 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGTCTGGTCAGGTCGCGGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA 360

Qy   121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCCAATCACCGGTACATACCCAGGCAGATACG 420

Qy   141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTGCCAACTGGATAACGCG 480

Qy   161 AsnValThrAspAlaIleLeuSerArgAlaAlaGlyGlySerIleLeuAlaAspPheThrGlyHis 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AACGTGACTGACGCGATCCTCAGCAGGGCCAGGAGGTCAATTGCTGACTTTACCGGGCAT 540

Qy   181 ArgGlnThrArgAlaPheArgGluLeuGlnLysArgLeuValLeuAsnPheProGlnSerLeuAsnCys 200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CGGCAAACGGCGTTTCGCGAACTGGAAACGCGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600

Qy   201 LeuLysArgGluLysGlnAspAspGluSerCysSerSerLeuThrGlnAlaAlaLeuProSerGluLeu 220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCAATTACCATCGGAACTC 660

Qy   221 LysValSerAlaAspAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerLeuAlaSerMetLeuThr 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 720

Qy   241 GluIlePheLeuLeuGlnLeuGlnAlaGlnLnGlyMetProGlyProGlyTrpProGlyLysArgIleThr 260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GAGATATTTCTTCCTGCAACAGACAGCAGGGCCGGAATGCCCGGAGCCGGGGGAAGGATCACC 780

Qy   261 AspSerHisGlnLnTrpAsnThrLeuAsnLeuSerLeuHisAsnAlaGlnLnPheTyrLeuLeuGln 280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GATTCACACCAGTGGAACACACTTGCTAAGTTTGCATAACGGCAATTTTATTTGCTACAA 840

Qy   281 ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuLeuAspLeuIleMetAlaAlaAla 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CGCACGCCAGAGGTTGCCCAGCCGCGCCACCCCGTTATTGGATTTGAATCATGGCAGCG 900

Qy   301 LeuThrProHisProProGlnLysGlnAlaTyrArgValThrLeuProThrThrSerValLeu 320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTGACCCCATCCACCCGCAAAAACAGGGCTATGGTGTGACATTACCCACTTCAGTACTG 960

Qy   321 PheIleAlaGlyHisHisAspThrAsnLeuValAlaAsnLeuGlyLeuAlaLeuGlyLeuAsnTrp 340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TTTATTGCCGGACACCGATACTAATCTGGCAAATCTCGGCGGCCACTGGAGCTCAACTGG 1020

Qy   341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlyHisThrGlnGluLeuValPheGluGlnArgTrp 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

REFERENCE   1 (bases 1 to 1323)
AUTHORS   Kretz, K.
TITLE   Phytase
JOURNAL   Patent: US 6190897-A 1 20-FEB-2001;
FEATURES   Location/Qualifiers
  source   1..1323
                /organism="unknown"
                /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:   Length: 1323
Score: 8.92e-205   Matches: 440
Percent Similarity: 2302.00   Conservative: 0
Best Local Similarity: 100.0%   Mismatches: 0
Query Match: 100.0%   Indels: 0
DB: 2   Gaps: 0

US-10-601-319-2 (1-440) x AR130956 (1-1323)

```
Qy    1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProGlnSerAla       20
        | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    1 ATGAAGGCGATCTTAATCCCATTTTTATCTTCTGATTCCGTTAACCCGCAATCTGCA   60

Qy   21 PheAlaGlnSerGluProGluLeuLysValGluSerValValIleValSerArgHisGly   40
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   61 TTCGCTCAGAGTGAGCCCGAGCTGAAGGTGTGGAAAGTGTGTGATTGTCAGTCGTCATGGT  120

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro   60
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA  180

Qy   61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu   80
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC  240

Qy   81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro      100
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  241 GGACATTACCAACGCCAGCGCCTCGTGGCAGATGGTCTGCGGCGAAAAAGGGCGTGCCCG  300

Qy  101 GlnSerGlyGlnValAlaIleIleAlaAspGluThrValHisThrGlnAspThr         120
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  301 CAGTCTGGTCAGGTGCAGGTCGCGATTATTGCCGACGAGACGGTACCCGTAAAACAGGCGAA  360

Qy  121 AlaPheAlaAlaGlyLeuLeuAlaProAspCysAlaIleThrValHisThrGlnAspThr   140
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  361 GCCTTCGCCGCCGGGGCTGGCACCTGACTGTGCAATAACCAGGCAGATACG          420

Qy  141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla   160
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG  480

Qy  161 AsnValThrAspAlaIleLeuSerArgAlaGlyLeuSerIleAlaAspPheThrGlyHis   180
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  481 AACGTGACTGACGCGATCCTCAGCAGGGCAGGGGTCAATTGCTGACTTTACCGGGCAT     540

Qy  181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnProSerLeuProSerGluLeuCys   200
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  541 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC  600

Qy  201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu   220
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  601 CTTAAACTGAGAGAAACAGGACGAAAGCTGTTCATTACGGACCATTACCACCATGGAACTC  660

Qy  221 LysValSerAlaAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr   240
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  661 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG   720

Qy  241 GluIlePheLeuLeuGlnIleAlaGlnGlyMetProGlyLeuTrpTyrGlyTyrArgIleThr   260
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGGTCTGTGGTATGGGTAGGAGATCACC  780

Qy  261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln   280
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA  840

Qy  281 ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspIleMetAlaAla   300
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  841 CGCACGCCAGAGGTTGCCCGCCAGCCGCGCGCCACCCCGTTATTGGATATTGATCATGGCAGCG  900

Qy  301 LeuThrProHisProGlnLysProThrProLeuAsnAlaTyrGlyTyrValThrProThrSerValLeu   320
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  901 TTGACGCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTACTG   960

Qy  321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp   340
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  961 TTTATTGCCGGACACGATACTAATCTCGCAAATCTCGGCGGGGCACTGGAGCTCAACTGG  1020

Qy  341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyValGluLeuValPheGluArgTrp   360
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1021 ACGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGTGGGGTCGAACTGGTGTTTGAACGCTGG  1080

Qy  361 ArgGlyLeuSerAspAspAsnSerGlnTrpIleGlyGlnLeuSerLeuValPheGlnInThrLeuGln   380
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1081 CGTGGCCTTAAGCGATAACAGCCAGTGGATTGGTCAGTTTCGGTTTCCGCTCTGTCTTCCAGACTTTACAG  1140

Qy  381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr   400
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCGGGAGAGGTGAAAACTGACC  1200

Qy  401 LeuAlaGlyCysGluGluArgGlnAsnAlaGlnIleGlyMetCysSerLeuAlaGlyPheThrGln   420
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1201 CTGGCAGGATGTGAAGAGCGAAAATGCCAGGGCATGTGTTCGTTGGCAGGTTTTACCGCAA  1260

Qy  421 IleValAsnGluIleAlaArgIleProLeuAlaCysSerLeuArgSerHisHisHisHisHis   440
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1261 ATCGTTGAATGAGAGAAGCACGCATACCGGCCGTGCAGTTGAGATCTCATCACCATCAC   1320
```

RESULT 3
AR130956
LOCUS   AR130956   1323 bp   DNA   linear   PAT 16-MAY-2001
DEFINITION   Sequence 1 from patent US 6190897.
ACCESSION   AR130956
VERSION   AR130956.1 GI:14119281
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM   Unknown.
             Unclassified.

Db 781 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
         GATTCACCAGTGGAACACCTGCTAAGTTGCATAACGGCAATTTATTTGCTACAA   840

Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuMetAlaAla 300
Db 841 CGGACGCCAGAGGTTGCCCGCAGCCGCGCGCCACCCCGTTATTGGATTTGATCATGGCAGCG 900

Qy 301 LeuThrProHisProProLysGlnLysAlaTyrArgGlyValThrLeuProThrSerValLeu 320
Db 901 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGACATTACCACTTCAGTACTG 960

Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 961 TTTATTGCGGGACACGATACTAATCTGGCAAATCTCGGCGGCCACTGGAGCTCAACTGG 1020

Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyLeuValPheGluGluTrp 360
Db 1021 ACGCTTCCCGGTCAGCCGGATAACACCCCGCCCGGATTGGTGTTTGAACTGGAGTGG 1080

Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTTACAG 1140

Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProGlyValLysLeuThr 400
Db 1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACCCCGGGAGTGAAACTGACC 1200

Qy 401 LeuAlaGlyCysGluGluLeuArgIleLeuProAlaCysSerLeuArgSerHisHisHisHis 420
Db 1201 CTGGCAGGATGTGAAGAGCAGGGCATGTGTCGTTGCTGCAGGTTTTACGCAA 1260

Qy 421 IleValAsnGluArgIleLeuProAlaLeuArgIleProHisProHisHisHisHis 440
Db 1261 ATCGTGAATGAAGCGCACCACATACCGGCGTGCAGTTTGAGATCTTCATCACCATCAC 1320

RESULT 4
AR492966
LOCUS   AR492966   1323 bp   DNA   linear   PAT 15-MAY-2004
DEFINITION Sequence 1 from patent US 6720014.
ACCESSION   AR492966
VERSION   AR492966.1 GI:47264255
KEYWORDS   Unknown.
SOURCE   Unknown.
ORGANISM   Unclassified.
REFERENCE   1 (bases 1 to 1323)
AUTHORS   Short,J.M. and Kretz,K.A.
TITLE   Phytase-containing foodstuffs and methods of making and using them
JOURNAL   Patent: US 6720014-A 1 13-APR-2004;
            Diversa Corporation; San Diego, CA
FEATURES   Location/Qualifiers
source   1..1323
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:   8.92e-205   Length: 1323
Score: 2302.00   Matches: 440
Percent Similarity: 100.0%   Conservative: 0
Best Local Similarity: 100.0%   Mismatches: 0
Query Match: 100.0%   Indels: 0
DB: Gaps: 2

US-10-601-319-2 (1-440) x AR492966 (1-1323)

Qy   1 MetLysAlaIleLeuIlePheProLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db   1 ATGAAAGCGATCTTAATCCCATTTTCTCTGCTTATCCCGCTTAACCCCCAATCTGCA 60

Qy 21 PheAlaGlnSerGluProGluLeuLysProGluSerValValIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120

Qy 41 ValArgAlaProThrLysMetGlnAspValThrAspAlaTrpPro 60
Db 121 GTGCGTGCTCCAACCAAGGCCACGGATGTCACCGACGCATGGCCA 180

Qy 61 ThrTrpProValLysGlyTrpLeuThrProArgGlyGlyGluIleAlaTyrLeu 80
Db 181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGGAGCTAATCGCCTATCTC 240

Qy 81 GlyHisTyrArgGlnArgGlyLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
Db 241 GGACATTACCACCGCGCTGTCTAGCCGACGGATTGCTGGCAGAAAAGGCTGCCCG 300

Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCTGGTCAGGTCGCGGATTATTGCTGATGTCGACGAGCGGTACCCGTAAAAACAGGGGAA 360

Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleIleThrValHisThrValHisAspThr 140
Db 361 GCCTTCGCCGCGGGCTGGCACCTGACCTGTGCAATAACCGTACATACCCAGGCAGATACG 420

Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACNCGGCGTTTGCCAACTGGATAACGCG 480

Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACGGATGCTATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 540

Qy 181 ArgGlnThrAlaPheArgGluGluLeuValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 CGGCAAACGGCGTTTCGCGAACTGGAACGGTGCTTAATTTCCGAATCAAACTTGTGC 600

Qy 201 LeuLysArgGluValGlnAspLeuSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGCAGCATTACCATCGGAACTC 660

Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 720

Qy 241 GluIlePheLeuLeuGlnIleAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 721 GAGATATTTTCTCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGTGGGGAAGGATCACC 780

Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnInPheTyrLeuLeuGln 280
Db 781 GATTCACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA 840

Qy   401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA 1260

Qy   421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHisHis 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAGCAGCACGCATACCGGCGTGCAGTTTGAGATCTCATCACCATCACCAC 1320

RESULT 5
AR636185
LOCUS   AR636185   1323 bp   DNA   linear   PAT 20-APR-2005
DEFINITION   Sequence 1 from patent US 6855365.
ACCESSION   AR636185
VERSION   AR636185.1 GI:62768868
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE   1 (bases 1 to 1323)
  AUTHORS   Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B.,
             O'Donoghue,B. and Mathur,E.J.
  TITLE   Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: US 6855365-A 1 15-FEB-2005;
             Diversa Corporation; San Diego, CA
FEATURES   Location/Qualifiers
  source   1..1323
                  /organism="unknown"
                  /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:   8.92e-205   Length: 1323
Score: 2302.00   Matches: 440
Percent Similarity: 100.0%   Conservative: 0
Best Local Similarity: 100.0%   Mismatches: 0
Query Match: 100.0%   Indels: 0
DB: 2   Gaps: 0

US-10-601-319-2 (1-440) x AR636185 (1-1323)

Qy   1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla   20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA   60

Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly   40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTATTGTCAGTCGTCATGGT 120

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy   61 ThrTrpProValLysGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu   80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCCGCNGGTGGTGAGCTAATCGCCTATCTC 240

Qy   81 GlyHisTyrGlnArgGlnAsnProLeuValAlaAspProGlyLeuLeuAlaLysGlyCysPro 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGACATTACCAACGCCAGCGTCGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG 300

Qy   101 GlnSerGlyGlnValAlaIleIleAlaAspAlaAspAlaArgThrArgLysThrGlyGlu 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGTCTGGTCAGGTCGCGATTATTGCTGATGCTGCGACGAGCCGTACCGGTAAAACAGGCGAA 360

Qy   121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420

Qy   141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCCAGTCCGGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG 480

Qy   161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180

Db   481 AACGTGACTGACGCGATCCTGAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 540

Qy   181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTCCCGCAATCAAACTTGTGC 600

Qy   201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC 660

Qy   221 LysValSerAlaAspAsnAlaSerLeuThrGlyAlaValAlaValSerLeuAlaSerMetLeuThr 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AAGGTGAGCGCCGACAATGTCTATTAACCGGTCGGTAAGCCTTCGCATCAATGCTGACG 720

Qy   241 GluIlePheLeuGluGlnGlnIleGlnAlaGlnAspGlnGlyMetProGluProGlyTrpGlyTyrArgIleThr 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GAGATATTTCTCCTGCAACAAGCACAGGAATGCCGGAGCCGGGGTGGGGAAGGATCACC 780

Qy   261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GATTCACACAGTGGAACACCTTGCTAAGTTGCATAACGCAATTTTATTGCTACAA 840

Qy   281 ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CGCACGCCAGAGGTTGCCGCGCCACCCCGTTATTGGATTTGATTGATCATGGCAGCG 900

Qy   301 LeuThrProHisProProGlnLysGluGlnAlaTyrGlyValThrArgGlyValThrSerValLeu 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTGACGCCCCATCCACCGCAAAAACAGGCTATGGTGCATTACCCACTTCAGTACTG 960

Qy   321 PheIleAlaGlyHisAspAsnAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TTTATTGCCGGACACGATAACAATCTGGCAAATCTCGGGGGCGCACTGGAGCTCAACTGG 1020

Qy   341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGCTTCCCGGTCAGCCGGATAACACCGCCGCAGGGTGGTGAACTGGTGTTTGAACGCTGG 1080

Qy   361 ArgArgLeuSerAspAsnSerGlnThrIleGlnValSerLeuValPheGlnThrLeuGln 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTCGCTGGTCTTCCAGACTTTACAG 1140

Qy   381 GlnMetArgAspProLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAGATGCGTGATAAAACCGCGCTGTCATTAAATACGCCCCCCGGAGAGGTGAAACTGACC 1200

Qy   401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGCAGGATGTGAAGAGCGAAATGCCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA 1260

Qy   421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHisHis 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAGCAGCACGCATACCGGCGTGCAGTTTGAGATCTCATCACCATCACCAC 1320

RESULT 6
AX052713
LOCUS   AX052713   1323 bp   DNA   linear   PAT 12-JAN-2001
DEFINITION   Sequence 1 from Patent WO0071728.
ACCESSION   AX052713
VERSION   AX052713.1 GI:12226903
KEYWORDS   .
SOURCE   Escherichia coli
  ORGANISM   Escherichia coli
             Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
             Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Short,J.M. and Kretz,K.A.
  TITLE   Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: WO 0071728-A 1 30-NOV-2000;
             Diversa Corporation (US)
FEATURES   Location/Qualifiers
  source   1..1323

```
                    221  LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerAlaSerMetLeuThr  240
                Qy       |||||||||||||||||||||||||||||||||||||||||||||||||||||||
                    661  AAGGTAGCGGACAATGTCTCATTAACCGGCGTGGGTAAGCCTCGCATCAATGCTGACG  720
                Db

                    241  GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGlyTrpProGlyTyrArgIleThr  260
                Qy       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                    721  GAGATATTTTCCTGCAACAAGCACAGGAGGAATGCCGGAGCCGGGGTCGGGAAGGATCACC  780
                Db

                    261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln  280
                Qy       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                    781  GATTCACCACCAGTGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA  840
                Db

                    281  ArgThrProGluValAlaSerArgAlaThrProLeuAspLeuIleMetAlaAla  300
                Qy       |||||||||||||||||||||||||||||||||||||||||||||||||||
                    841  CGCACGCCAGAGGTTGCCGCAGCCGCGCCACCCCGTTATTGGATTTGATCATGGCAGCG  900
                Db

                    301  LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu  320
                Qy       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                    901  TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACCTTCAGTACTG  960
                Db

                    321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340
                Qy       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                    961  TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG  1020
                Db

                    341  ThrLeuProGlyGlnProAspAsnThrProAsnThrProGlyAsnThrProGlyAsnArgTrp  360
                Qy       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                    1021 ACGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGTGGTGAACTCGGCTGG  1080
                Db

                    361  ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln  380
                Qy       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                    1081 CGTCGGCTAAGCAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG  1140
                Db

                    381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyValValGlyLeuThr  400
                Qy       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                    1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCCCCGAGAGGTGAAACTGACC  1200
                Db

                    401  LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln  420
                Qy       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                    1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA  1260
                Db

                    421  IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHisHis  440
                Qy       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                    1261 ATCGTGAATGAAGCACGCATACCGGCCTGCAGTTTGAGATCTCATCACCATCACCATCAC  1320
                Db
```

```
RESULT 7
AX338966
LOCUS            AX338966               1323 bp    DNA            linear    PAT 09-JAN-2002
DEFINITION       Sequence 1 from Patent WO0189317.
ACCESSION        AX338966
VERSION          AX338966.1  GI:18129102
KEYWORDS         .
SOURCE           Escherichia coli
  ORGANISM       Escherichia coli
                 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
                 Enterobacteriaceae; Escherichia.
REFERENCE        1
  AUTHORS        Short,J.M., Kretz,K.A. and O'Donoghue,E.
  TITLE          Dietary aids and methods of use thereof
  JOURNAL        Patent: WO 0189317-A 1 29-NOV-2001;
                 DIVERSA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..1323
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
     CDS             1..1323
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAD20553.1"
                     /db_xref="GI:18129103"
                     /translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAP
                     TKATQLMQDVTPDAWPTWVKLGWLTPRGELIAYLGHYQRQRLVADGLLAKKGCPQS
```

```
                CDS              /organism="Escherichia coli"
                                 /mol_type="unassigned DNA"
                                 /db_xref="taxon:562"
                                 1..1323
                                 /note="unnamed protein product"
                                 /codon_start=1
                                 /transl_table=11
                                 /protein_id="GI:12226904"
                                 /db_xref="GI:12226904"
                                 /translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAP
                                 TKATQLMQDVTPDAWPTWVKLGWLTPRGELIAYLGHYQRQRLVADGLLAKKGCPQS
                                 GQVAIIADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA
                                 NVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPS
                                 ELKVSADNVSLTGAVSLASMLTEIFLLQQAGMPEPGWGRITDSHQWNTLLSLHNAQF
                                 YLIQRTPEVARSRATPLLDLIMAALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGG
                                 ALELNWTLPGQPDNTPPGGBLVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNT
                                 PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSLRSHHHHHH"

ORIGIN

Alignment Scores:
Pred. No.:                        8.92e-205      Length:              1323
Score:                            2302.00        Matches:              440
Percent Similarity:               100.0%         Conservative:           0
Best Local Similarity:            100.0%         Mismatches:             0
Query Match:                      100.0%         Indels:                 0
DB:                                              Gaps:                   2
```

```
US-10-601-319-2 (1-440) x AX052713 (1-1323)

                    1   MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla  20
                Qy      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                    1   ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA  60
                Db

                    21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly  40
                Qy      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                    61  TTCGCTCAGAGTGAGCCGGAGCTGAAACTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT  120
                Db

                    41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro  60
                Qy      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                    121 GTGCGGGTGCTCCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCGGCCA  180
                Db

                    61  ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGluLeuIleAlaTyrLeu  80
                Qy      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                    181 ACCTGGCCGTGGTAAAACTGGGTTGGCTGACACCGCGNGGTGAGCTAATCGCCTATCTC  240
                Db

                    81  GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro  100
                Qy      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                    241 GGACATTACCCACCGCGTCTGGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG  300
                Db

                    101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu  120
                Qy      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                    301 CAGTCTGGTCAGGTGGCGATTATTGCTGCGGATGTCGACGAGCGTACCCGTAAAACAGGCGAA  360
                Db

                    121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr  140
                Qy      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                    361 GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCCAATAACCGTACATACCCAGGCAGATACG  420
                Db

                    141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla  160
                Qy      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                    421 TCCAGTCCGGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG  480
                Db

                    161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis  180
                Qy      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                    481 AACGTGACTGACGCGGATCCTGCAGGGCGCAGGAGGGTCAATTGCTGACTTTTACCGGGCAT  540
                Db

                    181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys  200
                Qy      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                    541 CGGCAAACGGCCGTTCCGGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC  600
                Db

                    201 LeuLysLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220
                Qy      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                    601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC  660
                Db
```

```
              GQVAIIADVDBRTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA
              NVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCKREKQDESCSLTQALPS
              ELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQF
              YLLQRTPEVARSRATPLLDLIMAALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGG
              PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSLRSHHHHHH"

ORIGIN
Alignment Scores:
Pred. No.:         8.92e-205        Length:        1323
Score:             2302.00          Matches:       440
Percent Similarity: 100.0%          Conservative:  0
Best Local Similarity: 100.0%       Mismatches:    0
Query Match:       100.0%           Indels:        0
DB:                                 Gaps:          2

US-10-601-319-2 (1-440) x AX338966 (1-1323)

Qy     1  MetLysAlaIleLeuIleProPheLeuSerLeuIleLeuProLeuThrProGlnSerAla  20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  ATGAAAGCGATCTTAATCCCATTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA  60

Qy    21  PheAlaGlnSerGluProGluLysGluSerValValIleValSerArgHisGly        40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  TTCCCTCAGAGTGAGCCCGAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT       120

Qy    41  ValArgAlaProThrLysAlaThrGlnThrLeuMetGlnAspValThrProAspAlaTrpPro  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA  180

Qy    61  ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu  80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  ACCTGGCCGGTAAAACTGGTTGGCTGACACCCGCNGGTGGTGAGCTAATCGCCTATCTC  240

Qy    81  GlyHisTyrArgGlnArgLeuValAlaAlaAspGlyLeuLeuAlaLysLysGlyCysPro  100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  GGACATTACCAACCCAGCGTCTCGGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG  300

Qy   101  GlnSerGlyGlnGlnGlnValAlaAlaIleIleAlaAspValAspArgThrArgLysThrGlyGlu  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  CAGTCTCGTCAGGTCGGCATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA  360

Qy   121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr  140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  GCCTTCGCCGCCGGGGCTGGCCACCTGACCTGTGCAATAACCGTACCATACCCAGGCAGATACG  420

Qy   141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla  160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG  480

Qy   161  AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  AACGTGACTGACGCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT  540

Qy   181  ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys  200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  CGGCAAACGGCGTTCGCGGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC  600

Qy   201  LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATGGGAACTC  660

Qy   221  LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG  720

Qy   241  GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr  260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCCGAGCCCGGGTGGGGAAGGATCACC  780

Qy   261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln  280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA  840
```

```
Qy   281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTGGATTTGATCATGGCAGCG  900

Qy   301  LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu  320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTACTG  960

Qy   321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp    340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  TTTATTGCCGGACGATACTAATCTGGCAAATCTCGGGGCGCCACTGGAGCTCAACTGG  1020

Qy   341  ThrLeuProGlyGlnProAspAsnThrProGlyGlyGluValLysGluLeuValArgTrp  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  ACGCTTCCCGGTCAGCCGGATAACACGCCCGGCGGTGGTGAACTGGTGTTGAACGCTGG  1080

Qy   361  ArgArgLeuSerAspAsnSerGlnThrIleGlnValSerLeuValPheGlnGlnThrLeuGln  380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG  1140

Qy   381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProGlyGlyValLysGluLeuThr  400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  CAGATGCGTGATAAAACGCCGTTGTCATTAAATACGCCCGGAGAGGTGAAACTGACC  1200

Qy   401  LeuValAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGGCATGTGTCGTTGGCAGGTTTTACGCAA  1260

Qy   421  IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHisHis  440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  ATCGTGAATGAAGCAGCACCGCATACCGGCGGTGCAGTTTGAGATCTCCATCACCATCAC  1320
```

```
RESULT 8
AX356566
LOCUS       AX356566                 1323 bp    DNA        linear   PAT  06-FEB-2002
DEFINITION  Sequence 1 from Patent WO0190333.
ACCESSION   AX356566
VERSION     AX356566.1  GI:18621053
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B. and
            O'Donoghue,E.
  TITLE     Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: WO 0190333-A 1 29-NOV-2001;
            DIVERSA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..1323
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
     CDS             1..1323
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAD23025.1"
                     /db_xref="GI:18621054"
                     /translation="MKAILIPFLSLLLIPLTPQSAFAQSEPELKLESVVISRHGVRAP
                     TKATGLMQDVTPDAWPTWPVKLGWLTPRGGELIAYLGHYQRLVADGLLAKKGCPQS
                     GQVAIIADVDRTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA
                     NVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCKREKQDESCSLTQALPS
                     ELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQF
                     YLLQRTPEVARSRATPLLDLIMAALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGG
                     ALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQNIQVSLVFQTLQQMRDKTPLSLNT
                     PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSLRSHHHHHH"
ORIGIN
Alignment Scores:
Pred. No.:    8.92e-205    Length:    1323
```

```
Score:                    2302.00      Matches:              440
Percent Similarity:       100.0%       Conservative:           0
Best Local Similarity:    100.0%       Mismatches:             0
Query Match:                           Indels:                 0
DB:                                    Gaps:                   2

US-10-601-319-2 (1-440) x AX356566 (1-1323)

Qy      1  MetLysAlaIleLeuIlePheLeuSerLeuLeuIleProLeuThrProGlnSerAla  20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA  60

Qy     21  PheAlaGlnSerGluProGluLeuLysLeuLysGluSerValValIleValSerArgHisGly  40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGAAAGGTGTGAAAGTTGTCAGTCGTCATGGT  120

Qy     41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTGACCCCAGACGCATGGCCA  180

Qy     61  ThrTrpProValLysGlyGlyTrpLeuThrProArgGlyGlyGluLeuIleLeuAlaTyrLeu  80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  ACCTGGCCGGTAAAACTGGTTGGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC  240

Qy     81  GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro  100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG  300

Qy    101  GlnSerGlyGlnValAlaIleAlaIleLeuAlaAspAspThrArgLysThrGlyLysGlu  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA  360

Qy    121  AlaPheAlaAlaAlaGlyLeuLeuAlaProAspCysAlaIleThrGlnAlaThrGlnAlaAspThr  140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  GCCTTCGCCGCCGGGCTGCCACCTGACGTCGCAATAACGTACACATACCCAGGCAGATACG  420

Qy    141  SerSerProAspProGluPheAsnProLeuLysThrGlyValCysGlnLeuAspAspAsnAla  160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  TCCAGTCCGGATCCGTTATTTAATCCTCTAAAACTGGCGTTTTGCCAACTGGATAACGCG  480

Qy    161  AsnValThrArgAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  AACGTGACTCGACGCGATCCTCAGCAGGCAGGAGGGCTCAATTGCTGACTTTACCGGGCAT  540

Qy    181  ArgGlnThrArgAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys  200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAAACTTGTGC  600

Qy    201  LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC  660

Qy    221  LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCCGTAAGCCTCGGATCAATGCTGACG  720

Qy    241  GluIlePheLeuLeuGlnLeuGlnGlyLeuMetProGlyGlyTrpProGlyTyrArgIleThr  260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  GAGATATTTCTCCTGCAACAGCACAGGGAATGCCGGACGCCGGGGTGGGAAGGATCACC  780

Qy    261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln  280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA  840

Qy    281  ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  CGCACGCCAGAGGTTGCCGCGAGCCGCGCCACCCCCGTATTGGATTTGATCATGGCAGCG  900

Qy    301  LeuThrProHisProGlyGlnLysAlaTyrGlyValThrLeuProThrSerValLeu  320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  TTGACGCCCCATCCAGGGCAAAAACAGGCTATGGTGACATTACCCACTTCAGTACTG  960

Qy    321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    961  TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG  1020

Qy    341  ThrLeuProGlyGlnProAspAsnArgThrProGlyAsnThrProProGlyGluValPheGluArgTrp  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  ACGCTTCCCGGTCAGCCCGATAACACCCGGTAACACGGTGGTGTTGAACGGCTGG  1080

Qy    361  ArgArgLeuSerAspAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln  380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  CGTCGGCTAAGCGGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG  1140

Qy    381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr  400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  CAGATGCGTGATAAAACGCCCGCTGTCATTAATAACCGCCGCGGAGGTGAAACTGACC  1200

Qy    401  LeuAlaGlyCysGluGluValGlnArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  CTGGCAGGATGTGAAGAGGCGAAATGCCAGGGCATGTTCGTTGGCAGGTTTTACCGCAA  1260

Qy    421  IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis  440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261  ATCGTGAATGAAGCACGCATACCGGCCTGCAGTTTGAGATCTCATCACCATCAC  1320
```

```
RESULT 9
AR636190
LOCUS       AR636190               1901 bp    DNA        linear    PAT   20-APR-2005
DEFINITION  Sequence 7 from patent US 6855365.
ACCESSION   AR636190
VERSION     AR636190.1  GI:62768873
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1901)
  AUTHORS   Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B.,
            O'Donoghue,E. and Mathur,E.J.
  TITLE     Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: US 6855365-A 7 15-FEB-2005;
            Diversa Corporation; San Diego, CA
FEATURES             Location/Qualifiers
     source          1..1901
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              2.74e-198       Length:               1901
Score:                  2235.00         Matches:               430
Percent Similarity:     99.5%           Conservative:            0
Best Local Similarity:  99.5%           Mismatches:              2
Query Match:            97.1%           Indels:                  0
DB:                                     Gaps:                    0

US-10-601-319-2 (1-440) x AR636190 (1-1901)

Qy      1  MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla  20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188  ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA  247

Qy     21  PheAlaGlnSerGluProGluLeuLysLeuLysGluSerValValIleValSerArgHisGly  40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248  TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGAAAGTGTGGTGATTGTCAGTCGTCATGGT  307

Qy     41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308  GTGCGTGCTCCAACCAAGGCCACCGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA  367

Qy     61  ThrTrpProValLysGlyGlyTrpLeuThrProArgGlyTyrLeuGluIleLeuAlaTyrLeu  80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    368  ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC  427

Qy     81  GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro  100
```

DEFINITION   Sequence 7 from Patent WO0190333.
ACCESSION   AX356572
VERSION   AX356572.1 GI:18621059
KEYWORDS   .
SOURCE   Escherichia coli
  ORGANISM   Escherichia coli
             Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
             Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B. and
             O'Donoghue,E.
  TITLE   Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: WO 0190333-A 7 29-NOV-2001;
             DIVERSA CORPORATION (US)
FEATURES   Location/Qualifiers
     source   1..1901
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

Alignment Scores:
Pred. No.:   2.74e-198   Length: 1901
Score: 2235.00   Matches: 430
Percent Similarity: 99.5%   Conservative: 0
Best Local Similarity: 99.5%   Mismatches: 2
Query Match: 97.1%   Indels: 0
DB: 2   Gaps: 0

US-10-601-319-2 (1-440) x AX356572 (1-1901)

```
Qy     1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuProGlnSerAla  20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCGATTCCTTAACCCCGCAATCTGCA 247

Qy    21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly  40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 307

Qy    41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308 GTGCGTGCTCCAACCAAGGCCACCCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 367

Qy    61 ThrTrpProValLysGlyTyrLeuThrProArgGlyGluLeuLeuIleAlaTyrLeu  80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368 ACCTGGCCGGTAAAACTGGTTGGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC 427

Qy    81 GlyHisTyrGlnArgGlnArgLeuValAlaLeuAlaAspCysAlaAlaLysLysGlyCysPro 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGCTGCGGCGAAAAAGGGCTGCCCG 487

Qy   101 GlnSerGlyValGlnValAlaIleIleAlaProThrAspValSerThrGlnIleThrArgLeu 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488 CAGTCTGGTGTCCAGGTTCGCGATTATTGCTGACGTGTCGACGAGGTACCGGTAAAACAGGCGAA 547

Qy   121 AlaPheAlaAlaGlyLeuAlaProCysAspAlaThrValHisThrGlnAlaAspThr 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548 GCCTTCGCCGCCGGGCTGGCACCTGTGATGTCGACGTGCAATAACGTACATACCCAGGCAGATACG 607

Qy   141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   608 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG 667

Qy   161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaGlyPheThrGlyHis 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668 AACGTGACTGACGCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGGCTTTACCGGGCAT 727

Qy   181 ArgGlnThrAlaPheArgGluLeuArgValLeuAsnPheProGlnSerAsnLeuCys 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   728 CGGCAAACGGCGTTTCGCGAACTGAGAGTGCTTAATTTTCCGCAATCAAACTTGTGC 787

Qy   201 LeuLysArgGluLysGlnAspAspSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
```

```
Db    788  CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCCGGAACTC  847
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    221  LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr  240
Db    848  AAGGTGAGCGGCCGACAATGTCCATTAACCGGTGCGTAAGCCTCGCATCAATGCTGACG   907

Qy    241  GluIlePheLeuLeuGlnAlaGlnAlaGlyMetProGluProGlyTrpGlyArgIleThr  260
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908  GAGATATTTCTCCTGCAACAAGCACAGGAATGCCGGAGCCCGGGTGGGGAAGGATCACC   967

Qy    261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln  280
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968  GATTCACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA   1027

Qy    281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla  300
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028  CGCACGGCCAGAGGTTGCCCGCAGCCGCGCCACCCCCGTTATTAGATTTGATCAAGACAGCG  1087

Qy    301  LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu  320
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088  TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCTG  1147

Qy    321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluGluAsnTrp  340
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148  TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGGTCAACTGG  1207

Qy    341  ThrLeuProGlyGlnProAspAsnThrProProGlyGluLeuValPheGluArgTrp  360
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208  ACGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGTGGTGAACTGGTTTGAACGCTGG   1267

Qy    361  ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln  380
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268  CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG  1327

Qy    381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr  400
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328  CAGATGCGTGATAAAACCCCGCTGTCATTAAATACCGCCCCGAGAGGTGAAACTGACC  1387

Qy    401  LeuAlaGlyCysGluGlyArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln  420
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388  CTGGCAGGATGTGAAGAGCGAAATGCCCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA  1447

Qy    421  IleValAsnGluAlaArgIleProAlaCysSerLeu  432
                 ||||||||||||||||||||||||||||||||
Db   1448  ATCGTGAATGAAGCACGCATACCGGCCTGCAGTTTG  1483
```

```
RESULT 11
ECOAPPAA
LOCUS        ECOAPPAA              1901 bp    DNA        linear   BCT 04-APR-2002
DEFINITION   Escherichia coli periplasmic phosphoanhydride phosphohydrolase
             (appA) gene, complete cds.
ACCESSION    M58708
VERSION      M58708.1  GI:145283
KEYWORDS     .
SOURCE       Escherichia coli
  ORGANISM   Escherichia coli
             Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
             Enterobacteriaceae; Escherichia.
REFERENCE    1  (bases 1 to 1901)
  AUTHORS    Dassa,J., Marck,C. and Boquet,P.L.
  TITLE      The complete nucleotide sequence of the Escherichia coli gene appA
             reveals significant homology between pH 2.5 acid phosphatase and
             glucose-1-phosphatase
  JOURNAL    J. Bacteriol. 172 (9), 5497-5500 (1990)
   PUBMED    2168385
REFERENCE    2  (sites)
  AUTHORS    Ostanin,K., Harms,E.H., Stevis,P.E., Kuciel,R., Zhou,M.M. and Van
             Etten,R.L.
  TITLE      Overexpression, site-directed mutagenesis, and mechanism of
             Escherichia coli acid phosphatase
  JOURNAL    J. Biol. Chem. 267 (32), 22830-22836 (1992)
   PUBMED    1429631
REFERENCE    3  (sites)
  AUTHORS    Ostanin,K. and Van Etten,R.L.
  TITLE      Asp304 of Escherichia coli acid phosphatase is involved in leaving
             group protonation
  JOURNAL    J. Biol. Chem. 268 (28), 20778-20784 (1993)
   PUBMED    8407904
FEATURES             Location/Qualifiers
     source          1..1901
                     /organism="Escherichia coli"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:562"
     gene            3..1522
                     /gene="appA"
     RBS             3..8
                     /gene="appA"
     CDS             16..108
                     /gene="appA"
                     /note="ORFX"
                     /codon_start=1
                     /transl_table=11
                     /product="unknown"
                     /protein_id="AAA72085.1"
                     /db_xref="GI:145284"
                     /translation="MWYLLWFVGILLMCSLSTLVLVWLDPRLKS"
     stem_loop       118..149
                     /gene="appA"
     -35_signal      133..138
                     /gene="appA"
     -10_signal      156..161
                     /gene="appA"
     RBS             172..177
                     /gene="appA"
     CDS             188..1486
                     /gene="appA"
                     /standard_name="acid phosphatase"
                     /codon_start=1
                     /transl_table=11
                     /product="periplasmic phosphoanhydride phosphohydrolase"
                     /protein_id="AAA72086.1"
                     /db_xref="GI:145285"
                     /translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP
                     TKATQLMQDVTPDAWPTWPVKLGWLTPRGGELIAYLGHYORQRLVADGLLAKKGCPQS
                     GQVAIIADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNA
                     NVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPS
                     ELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQF
                     YLLQRTPEVARSRATPLLDLIKTALGTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGG
                     ALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNT
                     PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL"
     sig_peptide     188..253
                     /gene="appA"
     mat_peptide     254..1483
                     /gene="appA"
                     /product="periplasmic phosphoanhydride phosphohydrolase"
     variation       299..301
                     /gene="appA"
                     /standard_name="H303A mutant"
                     /note="created by site directed mutagenesis"
                     /citation=[2]
     variation       299..301
                     /gene="appA"
                     /standard_name="R16A mutant"
                     /note="created by site directed mutagenesis"
                     /citation=[2]
     variation       299..301
                     /gene="appA"
                     /standard_name="R20A mutant"
                     /note="created by site directed mutagenesis"
                     /citation=[2]
     variation       299..301
                     /gene="appA"
                     /standard_name="R92A mutant"
                     /note="created by site directed mutagenesis"
                     /citation=[2]
     variation       302..304
```

```
            /gene="appA"
            /standard_name="H17N mutant"
            /note="created by site directed mutagenesis"
            /citation=[2]
     1163..1165
            /gene="appA"
            /standard_name="D304A mutant"
            /note="created by site directed mutagenesis"
            /citation=[3]
            /phenotype="Asp 304 changed to Ala"
            /replace="gct"
     1163..1165
            /gene="appA"
            /standard_name="D304E mutant"
            /note="created by site directed mutagenesis"
            /citation=[3]
            /phenotype="Asp 304 changed to Glu"
            /replace="gag"
     1497..1522
            /gene="appA"
            /note="rho-independent terminator"

ORIGIN
Alignment Scores:
Pred. No.:            2.74e-198    Length:        1901
Score:               2235.00      Matches:       430
Percent Similarity:  99.5%        Conservative:  0
Best Local Similarity: 99.5%      Mismatches:    2
Query Match:         97.1%        Indels:        0
DB:                  15           Gaps:          0

US-10-601-319-2 (1-440) x ECOAPPA (1-1901)

Qy     1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla   20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA  247

Qy    21 PheAlaGlnSerGluProGluLeuLysGluSerValValIleValSerArgHisGly   40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248 TTCGCTCAGAGTGAGCCGGAGCTGAAGTGGGTAGCTGGATTGTCAGTCGTCATCGT  307

Qy    41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA  367

Qy    61 ThrTrpProValLysGlyTrpThrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu   80
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCCGCNGGGTGGTGAGCTAATCGCCTATCTC  427

Qy    81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysProPro  100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGCTCGCGAAAAAGGGCTGCCCG  487

Qy   101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA  547

Qy   121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr  140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548 GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG  607

Qy   141 SerSerProAspPheAsnProLeuLysThrGlyValCysGlnLeuAspLysAsnAla  160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   608 TCCAGTCCCGATCCGTTTAATTAATCCCTAAAAAACTGGCGTTGCCAACTGGATAACGCG  667

Qy   161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668 AACGTGACTGACGCGATCCTCAGCAGGGCCAGGAGGGTCAATTGCTGACTTTACCGGGCAT  727

Qy   181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys  200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   728 CGGCAAACGGCGTTTCGCGAACTGGAACGCGTGCTCAATTTTCCGCAATCAATCAAACTTGTGC  787
```

```
Qy   201 LeuLysArgGluValGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   788 CTTAAACGTGAGAAACAGGAGGAAAGCTGTTCATTAACCCAGGCATTACCATCGGAACTC  847

Qy   221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr  240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   848 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCCATCAATGCTGACG  907

Qy   241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr  260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   908 GAGATATTTTCTCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGTGGGAAGGATCACC  967

Qy   261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln  280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   968 GATTCACCACCAGTGGAACACCTTGCTAAGTTTGCATAACCGCCAATTTATTTGCTACAA  1027

Qy   281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1028 CGCACGCCAGAGTTGCCCGCGCCACCCCCGTTATTAGATTTGATCAAGACAGCG  1087

Qy   301 LeuThrProHisProProGlnLysGlnLysIleAlaTyrGlyValThrLeuProThrSerValLeu  320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1088 TTGACGCCCCATCCACCGCAAAAACAGGGTATGGTGTGACATTACCCACTTCAGTGCTG  1147

Qy   321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGluLeuAsnTrp  340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1148 TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGGCGGCCACTGGAGCTCAACTGG  1207

Qy   341 ThrLeuProGlyAsnThrProProGlyGlyGluLeuValPheGluArgTrp  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1208 ACGCTTCCCGGTCAGCCGGGATAACACGCCCCAGGTGGTGAACTGGTGTTTGAACGCTGG  1267

Qy   361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnGlnThrLeuGln  380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1268 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG  1327

Qy   381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyGluValLysLeuThr  400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1328 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCCGGAGAGGTGAAACTGACC  1387

Qy   401 LeuAlaGlyCysGluArgAsnAlaGlyLysMetCysSerLeuAlaGlyPheThrGlnGln  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1388 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA  1447

Qy   421 IleValAsnGluValAsnArgIleProAlaCysSerLeu  432
          |||||||||||||||||||||||||||||||||||||||||
Db  1448 ATCGTGAATGAAGCAGCACCGCATACCGGCGGTGCAGTTTG  1483
```

```
RESULT 12
AX042376
LOCUS        AX042376          3470 bp        DNA       linear    PAT 23-NOV-2000
DEFINITION   Sequence 5 from Patent WO0064247.
ACCESSION    AX042376
VERSION      AX042376.1  GI:11340994
KEYWORDS     .
SOURCE       synthetic construct
  ORGANISM   synthetic construct
             other sequences; artificial sequences.
REFERENCE    1
  AUTHORS    Forsberg,C.W., Golovan,S. and Phillips,J.P.
  TITLE      Transgenic animals expresing salivary proteins
  JOURNAL    Patent: WO 0064247-A 5 02-NOV-2000;
             UNIVERSITY OF GUELPH (CA)
FEATURES             Location/Qualifiers
     source          1..3470
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="R15/APPA transgene"
ORIGIN

Alignment Scores:
Pred. No.:            6.59e-198        Length:        3470
```

```
Db    2771  TTTATCGCGGACACGATACTAATCTGCAAATCTCCGCGGCGCCACTCGAGCTCAACTGG  2830
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     341  ThrLeuProGlyGlnProAspAsnSerGlnTrpIleGlnValSerLeuValPheArgTrp  360
Db    2831  ACGCTTCCCGGTCAGCCGGATAACACGCCGATAAACGCCGCCAGGTGGTGAACTGGTTGAAACGCTGG  2890
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     361  ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln  380
Db    2891  CGTCGGCTAAGCGAGCCAGTGGATGACAGCCAGTTCAGGTTTCGTCGTCTTCCAGACTTTACAG  2950
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProGlyGluValLysLeuThr  400
Db    2951  CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGGGCGAAGTGAAACTGACC  3010
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     401  LeuAlaGlyCysGluGluArgAsnAlaGlnAlaGlyPheThrGln  420
Db    3011  CTGGCAGGATGTGAAGAGCGAAATGCCCAGGCGCATGTGTTCGTTGGCAGGTTTTTACCGCAA  3070
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     421  IleValAsnGluAlaArgIleProAlaCysSerSerLeu  432
Db    3071  ATCGTGAATGAAGCAGCCAGCACGATACCCGCTTGCAGTTTG  3106
```

```
RESULT 13
AX042374
LOCUS        AX042374                4060 bp    DNA      linear    PAT 23-NOV-2000
DEFINITION   Sequence 3 from Patent WO0064247.
ACCESSION    AX042374
VERSION      AX042374.1  GI:11340992
KEYWORDS     .
SOURCE       synthetic construct
  ORGANISM   synthetic construct
             other sequences; artificial sequences.
REFERENCE    1
  AUTHORS    Forsberg,C.W., Golovan,S. and Phillips,J.P.
  TITLE      Transgenic animals expressing salivary proteins
  JOURNAL    Patent: WO 0064247-A 3 02-NOV-2000;
             UNIVERSITY OF GUELPH (CA)
FEATURES             Location/Qualifiers
     source          1..4060
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="R15/APPA + intron transgene"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                                     Length:            4060
Score:            8.28e-198                     Matches:           430
Percent Similarity:  99.5%                     Conservative:      0
Best Local Similarity:  99.5%                  Mismatches:        0
Query Match:      97.1%                         Indels:            2
DB:               2                             Gaps:              0

US-10-601-319-2  (1-440) x AX042374  (1-4060)
```

```
Qy      1  MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla  20
Db   1811  ATGAAAGCCATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTCA  1870
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     21  PheAlaGlnSerGluProGluLeuLysLeuValValIleValSerArgHisGly  40
Db   1871  TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAGTGTGGTGATTGTCAGTCGTCATGGT  1930
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     41  ValArgAlaProThrLysAlaProThrLysAlaProThrGlnMetGlnAspValThrProAlaTrpPro  60
Db   1931  GTGCGTGCTCCAACCAAGGCCACGACCAACTGATGCAGGATGTCACCCCAGGCCATGGCCA  1990
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61  ThrTrpProValLysLysProGlnTrpLeuThrProLeuThrProThrSerValLeu  80
Db   1991  ACCTGGCCGGTAAAACTGGGTTGGCTGACATTACCCCACTTCAGTGCTG  2050
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     81  GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyGlyLeuLeuAlaLysLysCysPro  100
```

```
Score:             2235.00                      Matches:           430
Percent Similarity:  99.5%                     Conservative:      0
Best Local Similarity:  99.5%                  Mismatches:        2
Query Match:      97.1%                         Indels:            0
DB:               2                             Gaps:              0

US-10-601-319-2  (1-440) x AX042376  (1-3470)
```

```
Qy      1  MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla  20
Db   1811  ATGAAAGCCATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA  1870
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     21  PheAlaGlnSerGluProGluLeuLysLeuValValIleValSerArgHisGly  40
Db   1871  TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT  1930
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     41  ValArgAlaProThrLysAlaThrGlnMetGlnAspValThrProAlaTrpPro  60
Db   1931  GTGCGTGCTCCAACCAAGGCCACGACCAACTGATGCAGGATGTCACCCCAGGCCATGGCCA  1990
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     61  ThrTrpProValLysLysProGlnTrpLeuThrProLeuThrProThrSerValLeu  80
Db   1991  ACCTGGCCGGTAAAAACTGGGTTGGCTGACATTACCCCACTTCAGTGCTG  2050
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     81  GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyGlyLeuLeuAlaLysLysCysPro  100
Db   2051  GGACATTACCAACGCCAGCCGTCTCAGGCGTTGCTGCCGACGGATTGCTGCCGGAAAAAGGGCTGCCCG  2110
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    101  GlnSerGlyGlnValAlaIleIleAlaAlaGluIleLeuAlaAspValAspLysThrArgLysThrGlyGlu  120
Db   2111  CAGTCTGGTCAGGTGGCGATTATTGCTGCGGAAATTCGACGAGCGTACCCGTAAAACAGGCGAA  2170
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    121  AlaPheGluAlaLaGlyLeuAlaProAspCysLysAlaIleThrValHisThrGlnAlaAspThr  140
Db   2171  GCCTTCGCCGCGCCGGGCCTGGCCACCTGACTGTCAATAACCGTACATACCCAGGCCAGATACG  2230
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAspAsnAla  160
Db   2231  TCCAGTCCGGATCCGCTGTTTATTTAATCCTCTAAAAACTGGCGTTGCCAACTGGATAACGCG  2290
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    161  AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis  180
Db   2291  AACGTGACTGACGCGGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGCAT  2350
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    181  ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys  200
Db   2351  CGGCAAACGGGCGTTTCCGCAACTGGAACGGGGTGCTTAATTTTCCGCAATCAAACTTGTGC  2410
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    201  LeuLysArgGluLeuGluSerCysSerSerLeuThrGlnAlaLeuProSerGluLeu  220
Db   2411  CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTACGAGGCATTACCATCGGAACTC  2470
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    221  LysValSerAlaAspAsnValSerLeuThrGlyAlaValAlaSerMetLeuThr  240
Db   2471  AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGTAAGCCTGGCCATCAATGCTGACG  2530
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241  GluIlePheLeuLeuGlnGlnGlnLeuGlyMetProGluProGlyTyrGlyTrpGlyTyrArgIleThr  260
Db   2531  GAGATATTTCTCCTGCAACAAGCACACAGGGAAATGCCCGAGCCCGGGTGGGGAAGGATCACC  2590
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln  280
Db   2591  GATTCACACAGTGGAACACCTTGCTAAGTTTGCATAACGGCGAATTTTATTTGCTACAA  2650
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla  300
Db   2651  CGCACGCCGGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTAGATTTGATCATGGCAGCG  2710
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    301  LeuThrProHisLysProGlnLysLysAlaTyrGlyValThrLeuProThrSerValLeu  320
Db   2711  TTGACGCCCATCCACCGCAAAAACAGGCGTATGGTGACATTACCCACTTCAGTGCTG  2770
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340
```

```
LOCUS        AX042375            6116 bp    DNA            linear   PAT 23-NOV-2000
DEFINITION   Sequence 4 from Patent WO0064247.
ACCESSION    AX042375
VERSION      AX042375.1  GI:11340993
KEYWORDS     .
SOURCE       synthetic construct
  ORGANISM   synthetic construct
             other sequences; artificial sequences.
REFERENCE    1
  AUTHORS    Forsberg,C.W., Golovan,S. and Phillips,J.P.
  TITLE      Transgenic animals expressing salivary proteins
  JOURNAL    Patent: WO 0064247-A 4 02-NOV-2000;
             UNIVERSITY OF GUELPH (CA)
FEATURES             Location/Qualifiers
     source          1..6116
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="R15/APPA plasmid with pBLCAT3 vector"

ORIGIN

Alignment Scores:
Pred. No.:                1.5le-197     Length:          6116
Score:                    2235.00       Matches:          430
Percent Similarity:       99.5%         Conservative:       0
Best Local Similarity:    99.5%         Mismatches:         2
Query Match:              97.1%         Indels:             0
DB:                                     Gaps:               0

US-10-601-319-2 (1-440) x AX042375 (1-6116)
```

```
Qy     1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1811 ATGAAAGCCATCTTAATCCCATTTTATCCTTCGATTCCTTGATCCGTTAACCCGCAATCTGCA 1870

Qy    21 PheAlaGlnSerGluProGluLeuLysGluLeuGluSerValValIleValSerArgHisGly 40
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1871 TTCGCTCAGAGTGAGCCGGAGCTGGAAGTTGGTGTGATTGTCAGTCGTCATGGT 1930

Qy    41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1931 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 1990

Qy    61 ThrTrpProValLysLeuGlyTrpLeuThrProGlnArgGlyGlyGluLeuIleAlaTyrLeu 80
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1991 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGGGTGGTGAGCTAATCGCCTATCTC 2050

Qy    81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysGlyCysPro 100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2051 GGACATTACCAACGCCAGCGTCTTGGTAGCCGACGGATTGCTGGCGAAAAGGGCTGCCCG 2110

Qy   101 GlnSerGlyLeuValAlaIleIleAlaIleAlaProAspValAspGluSerValLysThrGlyGlu 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2111 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAAACAGGCGAA 2170

Qy   121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2171 GCCTTCGCCGCCGGGCTGGCACCCGACTGTGCAATAACCGTACATACCCAGGCAGATACG 2230

Qy   141 SerSerProAspPheAsnProLeuLysThrGlyThrGlyValCysGlnLeuAspAsnAla 160
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2231 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTGCCAACTGGATAACCGCG 2290

Qy   161 AsnValThrAspAlaIleLeuSerArgAlaGlyLeuGlySerIleAlaAspPheThrGlyHis 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2291 AACGTGACTGACGCCATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 2350

Qy   181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2351 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 2410

Qy   201 LeuLysLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  2051 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGCTGGCGAAAAGGGCTGCCCG 2110

Qy   101 GlnSerGlyIleValAlaIleIleAlaIleAlaProAspValAspGluSerValLysThrGlu 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2111 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAAACAGGCGAA 2170

Qy   121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2171 GCCTTCGCCGCCGGGCTGGCACCCGACTGTGCAATAACCGTACATACCCAGGCAGATACG 2230

Qy   141 SerSerProAspPheAsnProLeuLysThrGlyThrGlyValCysGlnLeuAspAsnAla 160
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2231 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTGCCAACTGGATAACCGCG 2290

Qy   161 AsnValThrAspAlaIleLeuSerArgAlaGlyLeuGlySerIleAlaAspPheThrGlyHis 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2291 AACGTGACTGACGCCATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 2350

Qy   181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2351 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 2410

Qy   201 LeuLysLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
```

```
Db  2051 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGCTGGCGAAAAGGGCTGCCCG 2110

Qy   101 GlnSerGlyLeuValAlaIleLeuGluAlaAspProAspCysAlaIleThrValHisThrAlaAspThr 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2111 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAAACAGGCGAA 2170

Qy   121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2171 GCCTTCGCCGCCGGGCTGGCACCCGACTGTGCAATAACCGTACATACCCAGGCAGATACG 2230

Qy   141 SerSerProAspPheAsnProLeuLysThrGlyThrGlyValCysGlnLeuAspAsnAla 160
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2231 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTGCCAACTGGATAACCGCG 2290

Qy   161 GluIlePheLeuLeuLeuGlnLeuAsnAlaGlnGlnProGlyGlyTrpGlyTyrGlyArgIleThr 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2291 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCCGGGGTGGGGAAGGATCACC 2350

Qy   181 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2351 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA 2410

Qy   201 ArgThrProGluValAlaArgSerArgAlaThrAlaProLeuLeuAspLeuIleMetAlaAla 220
```

```
Db  2051 GGCAGCATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGCTGGCGAAAAGGGCTGCCCG 2110

Qy   101 GlnSerGlyLeuValAlaIleLeuGluAlaAspProAspThr 120
           |||||||||||||||||||||||||||||||||||||||||||
Db  2111 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAAACAGGCGAA 2170

Qy   121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2171 GCCTTCGCCGCCGGGCTGGCACCCGACTGTGCAATAACCGTACATACCCAGGCAGATACG 2230

Qy   141 LeuThrProHisProProGlnLysValGlyAlaTyrGlyValThrArgLeuProThrSerValLeu 160
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2231 TTGACGCCCCATCCACCGCAAAAAGCGGTATGGTGTGACATTACCCGACTTCAGTGCTG 2290

Qy   161 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGluLeuAsnTrp 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2291 TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCCACTCGAGCTCAACTGG 2350

Qy   181 ThrLeuProGlyGlnProProGlyGlyLeuGluLeuValPheGluLysThrTrp 200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2351 ACGCTTCCCGGTCAGCCGGATAACACGCCCGGTGGTGAACTGGGTGTTGAACGCTGG 2410

Qy   201 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 220
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2351 CGTCGGCGTAAGCGATAACAGCCAGTTGGATTCAGGTTTCGCTGCGTCTTCCAGACTTTACAG 2410

Qy   201 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 220
```

```
Qy   401 LeuAlaLeuGlyLysGluGlyLeuArgGlyAsnAlaGlnGlnLysGlyMetCysSerSerLeuAlaGlyLyPheThrGln 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3011 CTGGCAGGAGATGTGAAGAGCCAGGGCATGTGTCGTTCGTTGGCAGGTTTTTACGCAA 3070

Qy   421 IleValAsnGluLeuAlaArgIleProAlaCysSerLeu 432
           |||||||||||||||||||||||||||||||||||||||
Db  3071 ATCGTGAATGAAGGCATACCCGCTTGCAGTTTG 3106

RESULT 14
AX042375
```

```
Best Local Similarity:  99.5%                    Mismatches:  2
Query Match:            97.1%                     Indels:      0
DB:                        2                      Gaps:        0

US-10-601-319-2 (1-440) x AX042373 (1-6708)

Qy    1  MetLysAlaIleLeuIlePheSerLeuLeuIlePheLeuProThrProGlnSerAla       20
Db 1811  ATGAAAGCCATCTTAATCCCATTTATATCTCTTCTGATTCCGTTAACCCGCAATCTGCA    1870

Qy   21  PheAlaGlnSerGluProGluLysLysGluLeuGluSerValValIleValSerArgHisGly   40
Db 1871  TTTCGCTCAGAGTGAGCCGGAGCTGAAGAAGTGGTGATTGTCAGTCGTCATGGT         1930

Qy   41  ValArgAlaProThrLysAsnLeuMetGlnAsnValThrProAspAlaTrpPro           60
Db 1931  GTGCGTGCTCCAACCAAGGCCACGATGCAGAACTGACGGAAGTCACCCCAGACAGCCA     1990

Qy   61  ThrThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu  80
Db 1991  ACCTGGCCGGTAAAACTGGGTTGGCTGACACCACCGGGTGGTGAGCTAATCGCCTATCTC   2050

Qy   81  GlyHisTyrGlnArgGlnAsnArgLeuAlaAspSerGlyLeuLeuAlaLysLysGlyCysPro 100
Db 2051  GGACATTACCAACGCCAGTGCTCGGTAGCCGACGACGGATTGCTGGCGAAAAAGGGCTGCCCG 2110

Qy  101  GlnSerGlyGlnValAlaIleIleIleGluAlaAspGluArgThrArgLysThrGlyGlu   120
Db 2111  CAGTCTGGTCAGGTCGCGATTATTGCTGATGATCGACGAGCGTACCCGTAAAACAGGCCAA  2170

Qy  121  AlaPheAlaAlaGlyLeuValAlaProAspCysAlaIleThrArgAlaIleHisThrGlnAlaAspThr 140
Db 2171  GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATACCGTACATACCCAGGCAGGATACG  2230

Qy  141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnIleValLysAsnAla 160
Db 2231  TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAAACGCG  2290

Qy  161  AsnValThrAspAlaIleLeuSerArgAlaGlyIleGlySerIleAlaAspPheThrGlyHis 180
Db 2291  AACGTGACTGACGCGATCCTCAGCAGGCAGGAGGTCAATTGCTGACTTTACCGGGCAT    2350

Qy  181  ArgGlnThrAlaPheArgGluLeuGlyLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 2351  CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC  2410

Qy  201  LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu   220
Db 2411  CTTAAACGTGAGAAACAGGACGAAAGTCTGTTCATTAACGCAGGCATTACCATCGGAACTC  2470

Qy  221  LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr   240
Db 2471  AAGGTGAGCGGCCGACAATGTCTCATTAACCGGTGCGTAAGCCTCGCATCAATGCTGACG   2530

Qy  241  GluIlePheLeuLeuGlnIleAlaGlyLeuMetProGlyGlyTrpAlaArgIleThr     260
Db 2531  GAGATATTTCTCTGCGACAACAGCACAGGAAGCCCGGAGCCCGGGGTGGGGAAGGATCACC  2590

Qy  261  AspSerHisGlnTrpAsnThrLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln     280
Db 2591  GATTCACCAGTGACACCTTGCTAAGTTTGCATAACGGCGCAATTTTATTTGCTACAA     2650

Qy  281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuMetAlaAla     300
Db 2651  CGCACGCCAGAGGTTGCCCGAGCCGCGCCACCCCGTTATTAGATTTGATCAAGAGACGCG   2710

Qy  301  LeuThrProHisProGlyLysGlnAlaTyrGlyValThrLeuProThrSerValLeu     320
Db 2711  TTGACGCCCATCACCGGACAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCTG   2770

Qy  321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp     340
Db 2771  TTTATCGCCGGACACGATACTAATCTCGGCGCACTGGAGCTCAACTGG            2830

Qy  341  ThrLeuProGlyGlnAsnProAspAsnThrProProGlyGlyLeuValPheGluLeuArgTrp 360
Db 2831  ACGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGTGGTAACTGGTGTTGAACGGCTGG   2890

Qy  361  ArgArgLeuSerAspAsnSerGlnTrpIleGlnLeuSerLeuValPheGlnLeuThrLeuGln 380
Db 2891  CGTCGGCAGGATGTGAAGAGCGAAATGCCGGCATGTGTTCGTTGGCAGGTTTTACAG    2950

Qy  381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuValLeuThr 400
Db 2951  CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCAGGTGGAAACTGACC     3010

Qy  401  LeuAlaGlyCysGluLysArgAsnAlaGlnAsnGlyMetCysSerSerLeuAlaGlyPheThrGln 420
Db 3011  CTGGCAGGATGTGAAGAGCGAAATGCCAGGCATGTGTTCGTTGGCAGGTTTTACGGCAA   3070

Qy  421  IleValAsnGluAlaArgIleProAlaCysSerLeu                           432
Db 3071  ATCGTGAATGAAGCACGCATACCCGGCTTGCAGTTTG                         3106

RESULT 15
AX042373
LOCUS       AX042373    6708 bp    DNA      linear   PAT 23-NOV-2000
DEFINITION  Sequence 2 from Patent WO0064247.
ACCESSION   AX042373
VERSION     AX042373.1  GI:11340991
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Forsberg,C.W., Golovan,S. and Phillips,J.P.
  TITLE     Transgenic animals expressing salivary proteins
  JOURNAL   Patent: WO 0064247-A 2 02-NOV-2000;
            UNIVERSITY OF GUELPH (CA)
FEATURES    Location/Qualifiers
  source    1..6708
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="R15/APPA + intron plasmid with pBLCAT3 vector"
ORIGIN

Alignment Scores:
Pred. No.:                          Length:         6708
Score:            1.72e-197         Matches:         430
Percent Similarity:  2235.00        Conservative:      0
                       99.5%
```

```
Qy      341 ThrLeuProGlyGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2831 ACGCTTCCGGTCAGCCGGATAACACCGCCCAGGTGGTGAACTGGTGTTTGAACGCTGG 2890

Qy      361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2891 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 2950

Qy      381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyGluValLysLeuThr 400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2951 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCCGGAGAGGTGAAACTGACC 3010

Qy      401 LeuAlaGlyCysGluArgAsnAlaGlnMetCysSerLeuAlaGlyPheThrGln 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3011 CTGGCAGGATGTGAAGAGCGAAATGCCAGGGCCATGTTCGTTGGCAGGTTTTACGCAA 3070

Qy      421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
            ||||||||||||||||||||||||||||||||||
Db     3071 ATCGTGAATGAATGAAGCACGCCATACCCGCTTGCAGTTTG 3106
```

Search completed: June 14, 2006, 13:53:05
Job time : 5289.78 secs

THIS PAGE LEFT BLANK

OM protein - protein search, using sw model

```
              Copyright  (c) 1993 - 2006 Biocceleration Ltd.
              GenCore version 5.1.9
```

Run on: June 13, 2006, 10:22:24 ; Search time 25.7339 Seconds
(without alignments)
1645.118 Million cell updates/sec

```
Title:          US-10-601-319-2
Perfect score:  2302
Sequence:       1 MKAILIPFLSLLIPLTPQSA........IVNEARIPACSLRSHHHHH 440

Scoring table:  BLOSUM62
Gapop 10.0 , Gapext 0.5
```

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```
Database :    PIR_80:*
           1: pir1:*
           2: pir2:*
           3: pir3:*
           4: pir4:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	2235	97.1	432	2	B36733	acid phosphatase (
2	2210	96.0	434	2	H90770	phosphoanhydride p
3	2210	96.0	444	2	D85633	hypothetical prote
4	943.5	41.0	441	2	AC0201	acid phosphatase (
5	545	23.7	413	2	P90773	periplasmic glucos
6	545	23.7	413	2	B85636	periplasmic glucos
7	541	23.5	417	2	S25627	glucose-1-phosphat
8	539	23.4	413	2	JV0087	glucose-1-phosphat
9	526	22.8	414	2	E87316	periplasmic phosph
10	524	22.8	413	2	AG0632	glucose-1-phosphat
11	139.5	6.1	423	1	S06167	acid phosphatase (
12	133	5.8	416	2	T16058	hypothetical prote
13	131	5.7	423	2	A33395	acid phosphatase (
14	122.5	5.3	421	2	S14742	acid phosphatase (
15	117	5.1	386	2	JH0610	acid phosphatase (
16	113.5	4.9	344	2	B89130	protein F52E1.8 [1
17	109	4.7	2554	1	TVFF7L	kinase-related pro
18	108.5	4.7	10223	2	T30225	polyketide synthas
19	107.5	4.7	457	2	A56925	paired box transcr
20	107	4.6	450	2	A54429	paired box transcr
21	106.5	4.6	1048	2	BVECSC	exonuclease (EC 3.
22	105.5	4.6	459	2	S52250	paired box transcr
23	104.5	4.5	2535	2	AC0304	probable hemolysin
24	102	4.4	638	2	A29440	signal recognition
25	101.5	4.4	397	2	C81716	hypothetical prote
26	101.5	4.4	479	1	JN0890	acid phosphatase (
27	101.5	4.4	524	2	H75530	probable acid-CoA
28	101.5	4.4	1047	2	C85535	ATP-dependent dsDN
29	101.5	4.4	1047	2	G90684	ATP-dependent dsDN
30	101	4.4	1662	2	T18540	mofA protein precu
31	101	4.4	26926	2	I38344	titin, cardiac mus
32	100.5	4.4	369	2	JQ2278	hydroxymethylbilan
33	100.5	4.4	438	2	S64682	acid phosphatase (
34	100.5	4.4	478	2	S58315	major plasmid tran
35	100.5	4.4	1214	2	AG2897	conserved hypothet
36	100.5	4.4	1387	2	A97673	probable periplasm
37	100.5	4.4	6260	2	T30228	polyketide synthas
38	100	4.3	328	2	A83572	pyridoxal phosphat
39	100	4.3	1374	1	VCBB17	major capsid prote
40	100	4.3	4135	2	T42629	tenascin-X - bovin
41	99	4.3	636	2	G81983	pilin glycosylatio
42	98.5	4.3	425	2	D83186	hypothetical prote
43	98.5	4.3	1034	2	AB0551	exonuclease SbcC (
44	97.5	4.2	578	2	H75256	malate oxidoreduct
45	97	4.2	440	2	T32457	hypothetical prote

ALIGNMENTS

```
RESULT 1
B36733
acid phosphatase (EC 3.1.3.2) precursor - Escherichia coli (strain K-12)
N;Alternate names: phosphoanhydride phosphohydrolase, periplasmic; phytase 6
C;Species: Escherichia coli
C;Date: 19-Apr-1991 #sequence revision 19-Apr-1991 #text change 09-Jul-2004
C;Accession: B36733; S18018; B64839; A26534; S17960; S33278
R;Dassa, J.; Marck, C.; Boquet, P.L.
J; Bacteriol. 172, 497-500, 1990
A;Title: The complete nucleotide sequence of the Escherichia coli gene appA reveals sign
A;Reference number: A36733; MUID:90368616; PMID:2168385
A;Accession: B36733
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-432 <DAS>
A;Cross-references: UNIPROT:P07102; UNIPARC:UPI00004750A; GB:M58708; NID:g145283; PIDN
R;Greiner, R.; Jany, K.D.
Biol. Chem. Hoppe-Seyler 372, 664-665, 1991
A;Title: Characterization of a phytase from Escherichia coli.
A;Reference number: S18018
A;Accession: S18018
A;Molecule type: protein
A;Residues: 23-33 <GRB>
A;Cross-references: UNIPARC:UPI000179815
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: B64839
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-432 <BLAT>
A;Cross-references: UNIPARC:UPI00004750A; GB:AE000200; GB:U00096; NID:g2367111; PIDN:AA
R;Touati, E.; Danchin, A.
Biochimie 69, 215-221, 1987
A;Title: The structure of the promoter and amino terminal region of the pH 2.5 acid phos
A;Reference number: A26534; MUID:87271766; PMID:3038201
A;Accession: A26534
A;Molecule type: DNA
A;Residues: 1-50, 'NAGCHPRRMANLAG',65,'T',67-74,'DV',77-111,'S',<TOU>
A;Cross-references: UNIPARC:UPI00016EDE2; GB:X05471; NID:g40925; PIDN:CAA2931.1; PID:g
R;Dassa, J.; Fsihi, H.; Marck, C.; Dion, M.; Kieffer-Bontemps, M.; Boquet, P.L.
Mol. Gen. Genet. 229, 341-352, 1991
A;Title: A new oxygen-regulated operon in Escherichia coli comprises the genes for a put
A;Reference number: S17958; MUID:92049231; PMID:1658595
A;Accession: S17960
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-17 <DA2>
A;Cross-references: UNIPARC:UPI00016FFB1; GB:S63811; NID:g238656; PIDN:AAB20286.1; PID:
```

R;Greiner, R.; Konietzny, U.; Jany, K.D.
Arch. Biochem. Biophys. 303, 107-113, 1993
A;Title: Purification and characterization of two phytases from Escherichia coli.
A;Reference number: S33278; MUID:93256556; PMID:8387749
A;Accession: S33278
A;Molecule type: protein
A;Residues: 23-31,'A',33-34 <GR2>
A;Cross-references: UNIPARC:UPI000179816
C;Comment: In addition to cAMP-mediated control, this enzyme is induced when bacterial c
naerobic conditions.
C;Genetics:
A;Gene: appA
A;Map position: 25 min
C;Keywords: monomer; periplasmic space; phosphohistidine; phosphoprotein; phosphoric mon
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-112/Product: acid phosphatase #status predicted <MAT>
F;38/Active site: Arg #status predicted
F;39/Active site: His (phosphohistidine intermediate) #status predicted

Query Match   97.1%; Score 2235; DB 2; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.2e-171;
Matches 430; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

```
Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCERNAQGMCSLAGFTQ  420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCERNAQGMCSLAGFTQ  420

Qy  421 IVNEARIPACSL 432
        ||||||||||||
Db  421 IVNEARIPACSL 432
```

RESULT 2
H90770
phosphoanhydride phosphorylase [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: H90770
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: H90770
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-434 <HAY>
A;Cross-references: UNIPROT:Q8XC29; UNIPARC:UPI0000D0517; GB:BA000007; PIDN:BAB34559.1;
A;Experimental source: strain O157:H7, substrain RIMD 0509952

C;Genetics:
A;Gene: ECs1136

Query Match   96.0%; Score 2210; DB 2; Length 434;
Best Local Similarity 98.4%; Pred. No. 1.2e-169;
Matches 427; Conservative   0; Mismatches   5; Indels   2; Gaps   1;

```
Qy    1 MKAILIPFLSLLIPLTPQSAFAQS--EPELKLESVVIVSRHGVRAPTKATQLMQDVTPDA  58
        ||||||||||||||||||||||||  |||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDA  60

Qy   59 WPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKT 118
        ||| ||||||||||||||||||||||||||||||||| |||||| |||||||||||||||
Db   61 WPNWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLTKKGCPOPGQVAIIADVDERTRKT 120

Qy  119 GEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFT 178
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFT 180

Qy  179 GHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASM 238
        |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db  181 GHRQTAFRELERVLNFPQSNLCLNRBKQDESCSLTQALPSELKVSADNVSLTGAVSLASM 240

Qy  239 LTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIM 298
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 LTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIM 300

Qy  299 AALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFE 358
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 IALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFE 360

Qy  359 RWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCERNAQGMCSLAGF  418
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCERNAQGMCSLAGF  420

Qy  419 TQIVNEARIPACSL 432
        ||||||||||||||
Db  421 TQIVNEARIPACSL 434
```

RESULT 3
D85633
hypothetical protein appA [imported] - Escherichia coli (strain O157:H7, substrain EDL93:
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: D85633
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D85633
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-444 <STO>
A;Cross-references: UNIPROT:Q8XC29; UNIPARC:UPI000165751; GB:AE005174; NID:g12514245; PI
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: appA

Query Match   96.0%; Score 2210; DB 2; Length 444;
Best Local Similarity 98.4%; Pred. No. 1.2e-169;
Matches 427; Conservative   0; Mismatches   5; Indels   2; Gaps   1;

```
Qy    1 MKAILIPFLSLLIPLTPQSAFAQS--EPELKLESVVIVSRHGVRAPTKATQLMQDVTPDA  58
        ||||||||||||||||||||||||  |||||||||||||||||||||||||||||||||
Db   11 MKAILIPFLSLLIPLTPQSAFAQSEPEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDA  70

Qy   59 WPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKT 118
        ||| ||||||||||||||||||||||||||||||||| |||||| |||||||||||||||
Db   71 WPNWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLTKKGCPOPGQVAIIADVDERTRKT 130

Qy  119 GEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFT 178
```

```
Db          ||||||||||||||||||||||||||||||||||||
    131 GEAPAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFT 190
Qy  179 GHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASM 238
        ||||||||||||||   |||  ||||||||||||||||||||||||||||||||||||
Db  191 GHRQTAFRELERVLNFPQSNLCLNREKQDESCSLTQALPSELKVSADNVSLTGAVSLASM 250
Qy  239 LTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIM 298
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  251 LTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIM 310
Qy  299 AALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFB 358
        ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  311 IALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFB 370
Qy  359 RWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGF 418
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  371 RWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGF 430
Qy  419 TQIVNEARIPACSL 432
        ||||||||||||||
Db  431 TQIVNEARIPACSL 444
```

RESULT 4
AC0201
acid phosphatase (EC 3.1.3.2) [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AC0201
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AC0201
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-441 <KUR>
A;Cross-references: UNIPROT:Q8ZFP6; UNIPARC:UPI00000DCCF7; GB:AL590842; PIDN:CAC90470.1;
C;Genetics:
A;Gene: YPO1648
C;Keywords: phosphoric monoester hydrolase

Query Match   41.0%;   Score 943.5; DB 2; Length 441;
Best Local Similarity 45.6%; Pred. No. 6.5e-68;
Matches 199; Conservative 66; Mismatches 156; Indels 15; Gaps 7;

```
Qy   5 LIPFLSLLIPLTPQSAFAQSEPE-LKLESVVIVSRHGVRAPTKATQLMQDVTPDAWPTWP 63
         ||:|:|   | |:|||||||||||   ||:|||:| :||:::| :|||  |||||||||
Db  13 LVLMLSGLAAITAPVA---AEPSGYTLERVVILSRHGVRSPTKQTQLMNDVTPDKNPQWP 69
Qy  64 VKLGWLTPRGGBLIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGEAFA 123
         |:|:|| ||||  ||   |   | :::   ||  | | |:||  || ||||||||||:||
Db  70 VKAGYLTPRGABLVTLMGGFYGDYFRSLGLIA-AGCPAEGGVYAQADIDQRTRLTCQAFL 128
Qy 124 AGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGHRQT 183
         ||:|||||||||||||||||||||||||||||||||||||||||| ::|
Db 129 DGVAPGCGLTVHNQADLKKTDPLFHPVEAGVCKLDAAQTDKAIEEQLGGPLDTVSQRYAK 188
Qy 184 AFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSAD--NVSLTGAVSLASMLTE 241
         |||||||||||||||||||||||||||||||||||||| :   ||:||||| :|||:|
Db 189 PFAQMGDVLNFAASPYCKSLQQQGKTCDFAHFAANEVNVNKEGTKVTLSGPLALSSTIGE 248
Qy 242 IFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAAL 301
         |||||||||||||| |||  | |:||:|||||||| ::|||||||||||||||||||
Db 249 IFLLQNAQAMPEVAWQRLKGAENWVSLLSLHNAQFNLMAKTPYIARHKGTPLLQQIDTAL 308
Qy 302 TPHPPQKQAYGVTLPTS---VLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVF 357
         | || |:|| |||||    ||: | |:|| |  ||| |:  ||||  |||||||||| |
Db 309 T---LQLDAQGQKLPISAQNRVLFLGGHDTNIANIAGMLGADWQLPEQPDNTPPGGGLVF 365
```

```
Qy  358 ERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSL-NTPPGEVKLTLAGCEERNAQGMCSLA 416
            :||  :  :: | |  |:|  | ::||: ::||::| |:|| :||| :::||||
Db  366 ELWQNPDNHQRYVAVKMFYQTMDQLRNAEKLDLKNNPAGIISVAVAGCENNGDDKLCELD 425
Qy  417 GFTQIVNEARIPACSL 432
            ||:| |||||||||
Db  426 TFQKKVAKVIEPACHI 441
```

RESULT 5
F90773
periplasmic glucose-1-phosphatase [imported] - Escherichia coli (strain O157:H7, substra
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: F90773
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A;Reference number: A99629; MUID:21155231; PMID:11258796
A;Accession: F90773
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-413 <HAY>
A;Cross-references: UNIPROT:Q8XBZ6; UNIPARC:UPI0O000D0533; GB:BA000007; PIDN:BAB34581.1;
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs1158

Query Match   23.7%;   Score 545; DB 2; Length 413;
Best Local Similarity 32.1%; Pred. No. 6.1e-36;
Matches 135; Conservative 76; Mismatches 180; Indels 30; Gaps 10;

```
Qy  13 IPLTPQSAPAQSEPB-LKLESVVIVSRHGVRAP-TKATQLMQDVTPDAWPTWPVKLGWLT 70
          :||| : ::  |  || ::  :|:::||  ||   |||:|| ::| ::| | ||:|
Db  13 VVLLASNAQAQTVPEGYQLQQVLMMSRHNLRAPLANNGSVLEQSTPNKWPEWDVPGGQLT 72
Qy  71 PRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGEAFAAGLAPDC 130
          |:|| :|| |::   |   | ::  | | ::    |  || |||||||  : |  ||
Db  73 TKGGVLEVVMGHYMREWLABQGMVKSGECPPPDTVYAYANSLQRTVATAQFFITGAFPGC 132
Qy 131 AITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGHRQTAFRELER 190
          :|  ||::  |:|||||| || |:|| |  |   ||:: |||:||  |  | ::|| |
Db 133 DIPVHHQEKMGTMDPTFNPVITD---DSAAFSEQAVAAMEKELSKL--QLTDSYQLLEK 186
Qy 191 VLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLTEIFLLQQAQG 250
          :|:  |||| | :   :|| ||  | | : || ::| :| |:|  :  |||||||||
Db 187 IVNYKDSPAC---KEKQQCSLVDGKNTFSAKYQQEPGVSGPLKVGNSLVDAFTLQYYEG 242
Qy 251 MP--EPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAALTPHPPQK 308
          |    |||||||||||||||:||||||| :||:|||||:||||:|| ||::
Db 243 FPMDQVAWGEIKSDQQWKVLSKLKNGYQDSLFTSPEVARNVAKPLVSYIDKAL------ 295
Qy 309 QAYGVTLPTS--VLFIAGHDTNLANLGGALELN-WTLPGQPDNTPPGGELVFERWRRLS 364
          ::  |||||    |:  ||:||| ||:  :| :  || ||:|||||| |:| :||| |
Db 296 ----VTDRTSAPKITVLVGHDSNIASLLTALDFKPYQLHDQNERTPIGGKIVFQRWRDSK 351
Qy 365 DNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGTQIVNE 424
          ||  |  :| :  | : |:|  || :  |  : |  |  |  :|   |  : |  | |
Db 352 ANRDLMKIEYVVQSAEQLRNADALTLQAPAQRVTLELSGC-PIDADGFCPMDKFDSVLNE 410
Qy 425 A 425
        |
Db 411 A 411
```

RESULT 6
B85636
periplasmic glucose-1-phosphatase [imported] - Escherichia coli (strain O157:H7, substra
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: B85636
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: B85636
A;Accession: B85636   MUID:21074935; PMID:11206551
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-413 <STO>
A;Cross-references: UNIPROT:Q8XBZ6; UNIPARC:UPI00000D0533; GB:AB005174; NID:g12514273; E
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: agp

Query Match   23.7%; Score 545; DB 2; Length 413;
Best Local Similarity   32.1%; Pred. No. 6.1e-36;
Matches 135; Conservative 76; Mismatches 180; Indels 30; Gaps 10;

Qy   13 IPLTPQSAFAQSEPE-LKLESVVIVSRHGVRAP-TKATQLMQDVTPDAWPTWPVKLGWLT 70
          |||| : :||: :||| |||:|:|||||||||| ||||||||||||||||||||||||:
Db   13 VVLLASNAQAQTVPEGYQLQQVLMMSRHNLRAPLANNGSVLEQSTPNKWPEWDVPGGQLT 72

Qy   71 PRGGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGEAFAAGLAPDC 130
          ||| :||:|||||||||| :|:| :|:::||||||||||||||||||:||:
Db   73 TKGGVLEVYMGHYMREWLAEQGMVKSGBECPPPDTVYAYANSLQRTVATAQFFITGAFPGC 132

Qy   131 AITVHTQADTSSPDDLFNPLKTGVCVQLDNANVTDAILSRAGGSIADFTGHRQTAFRELER 190
          :|: :: |||| :|| :::: : ||:||||||||||||:||:
Db   133 DIPVHHQEKMGTMDPFNFPVVTD----DSAAFSEQAVAAMEKELSKL--QLTDSYOLLEK 186

Qy   191 VLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLTEIFLLQQAQG 250
          :|: ||| : ::|:||:|||||||||||||||||||||||||||||||||||
Db   187 MP--EPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAALTPH 242

Qy   251 MP--EPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAALTPHPPQK 308
          |||:|||::|::|||||:::||||| || :|||||||||||||||||||||||||::
Db   243 FPMDQVAWGEIKSDQQWKVLSKLKNGYQDSLFTSPEVARNVAKPLVSYIDKAL------- 295

Qy   309 QAYGVTLPTS---VLFIAGHDTNLANLGGALEIN-WTLPGQPDNTPPGGELVFERWRRLS 364
          ||||||||| ||:|||||||||||||||||| ||||||||||||||||||||||:|
Db   296 ----VTDRTSAPKITVLVGHDSNIASLLTALDFKPYQLHDQNERTPIGGKIVFQRWRDSK 351

Qy   365 DNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQIVNE 424
          ||||||||||||||||||||||||||||||||:::| |||||||||||| :|:|
Db   352 ANRDLMKIEYVYYQSAEQLRNADALITLQAPAQRVTLELSGC-PIDADGFCPMDKFDSVLNE 410

Qy   425 A 425
          |
Db   411 A 411

RESULT 7
S25627
glucose-1-phosphatase (EC 3.1.3.10) precursor - Providencia rettgeri
C;Species: Providencia rettgeri
C;Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
A;Accession: S25627
R;Riccio, M.L.; Chiesurin, A.; Lombardi, G.; Satta, G.
submitted to the EMBL Data Library, September 1992
A;Reference number: S25627
A;Accession: S25627
A;Molecule type: DNA
A;Residues: 1-417 <RIC>
A;Cross-references: UNIPROT:Q52309; UNIPARC:UPI00001256FC; EMBL:X68201; NID:g45771; PIDN
A;Experimental source: strain PV7
C;Genetics:
A;Gene: agp
C;Keywords: periplasmic space; phosphohistidine; phosphoprotein; phosphoric monoester hy
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-417/Product: glucose-1-phosphatase #status predicted <MAT>
F;42/Active site: His (phosphohistidine intermediate) #status predicted
F;314/Active site: His #status predicted

Query Match   23.5%; Score 541; DB 2; Length 417;

Best Local Similarity   30.5%; Pred. No. 1.3e-35;
Matches 129; Conservative 84; Mismatches 182; Indels 28; Gaps 9;

Qy   10 SLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWPTWPVKLGWL 69
          :|:|: |: ||:|:|||| |:|||:|||||||||||:||||||||||||||||||||||
Db   13 ALFAPIAPTMASTDNQADMVLDQVLVLSRHNLRTPIVNTGILTEVTDKKWPDWDAKSGYL 72

Qy   70 TPRGGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQ-VAIIADVDERTRKTGEAFAAGLAP 128
          :||| :||:||||||||| : :|:|| ::   ::   :::||||||||||||||:||:
Db   73 TTQGGALEVYMGHYFRRWIDQNKLLADELCPTSNEDYLYTNSLQRTIATAQFPAAGAFP 132

Qy   129 DCAITVHTQADTSSPDDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTG--HRQTAFR 186
          ||:|:||::: : | :::   :   :::: ||||:||||||||||:
Db   133 GCKVNIHHQPBIGKMDPVFNPIIT-------NGSPEFKQKALAAMDDYLKGLSLKAGYE 184

Qy   187 ELERVLNFPPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLTEIFLLQ 246
          ||: ||| :| :|: ::|:|||
Db   185 ELDTVLNIKDSQKC----KTDKLCNLDSQKNSFIIEADKEPGVSGPLKIANSAVDAIDLQ 240

Qy   247 QAQGMP--EPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAALTPH 304
          ||:|: :||||||||:||||| :|||||||||||:|||||||||||||||||||||:
Db   241 YYEGFPADQVAWGLVDTPEKWCKLNTLKNAYQETLFTPKIIAKNVAHPILNYIDKGFSV 300

Qy   305 PPQKQAYGVTLPTSVLFIAGHDTNLANLGGALEIN-WTLPGQPDNTPPGGELVFERWRRL 363
          ||:| |||||||||||:|||||||||||||||||| ||||||||||||||||||||||:|
Db   301 DKGETA------KFIFLVGHDSNIASLMSAMDFKPYQLAQQYEHTPIGGKLVFQRWTDK 353

Qy   364 SDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCE-ERNAQGMCSLAGFTQIV 422
          ::|:|||||||||||||||||||||||||||:|||||||| |: |||||||||||:|:
Db   354 QTKKDFMKVEYVYYQTADQLRDNAYLSLETPPKHVTLELKDCPVDKN--GYCSWEDFQKVM 411

Qy   423 NEA 425
          :||
Db   412 AKA 414

RESULT 8
JV0087
glucose-1-phosphatase (EC 3.1.3.10) precursor - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
A;Accession: JV0087; H64841
R;Pradel, E.; Marck, C.; Boquet, P.L.
J. Bacteriol. 172, 802-807, 1990
A;Title: Nucleotide sequence and transcriptional analysis of the Escherichia coli agp ge
A;Reference number: JV0087; MUID:90130318; PMID:2153660
A;Accession: JV0087
A;Molecule type: DNA
A;Residues: 1-413 <PRA>
A;Cross-references: UNIPROT:P19926; UNIPARC:UPI00001256FB; GB:M33807; NID:g145217; PIDN:
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: H64841
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-413 <BLAT>
A;Cross-references: UNIPARC:UPI00001256FB; GB:AB000202; GB:U00096; NID:g1787233; PIDN:AA
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: agp
A;Map position: 23 min
C;Function:
A;Description: essential for growth in a high-phosphate medium containing glucose-1-phosp
A;Note: optimal at low pH
C;Keywords: homodimer; periplasmic space; phosphohistidine; phosphoprotein; phosphoric mo
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-413/Product: glucose-1-phosphatase #status predicted <MAT>
F;40/Active site: His (phosphohistidine intermediate) #status predicted
F;311/Active site: His #status predicted

```
Query Match                    23.4%;    Score 539;  DB 2;  Length 413;
Best Local Similarity   32.1%;    Pred. No. 1.8e-35;
Matches  135;  Conservative   75;  Mismatches  181;  Indels   30;  Gaps    10;

Qy     13  IPLTPQSAFAQSEPE-LKLESVVIVSRHGVRAP-TKATQLMQDVTPDAWPTWPVKLGWLT   70
               :::|:|::|||:  |:|:||||  ::||||||
Db     13  IVLLASNAQQTVPEGYQLQQVLMMSRHNLRAPLANNGSVLEQSTPNKWPEWDVPGGQLT   72

Qy     71  PRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGEAFAAGLAPDC  130
               ::|||||| |::        ::|
Db     73  TKKGVLEVVMGHYTMREWLAEQGMVKSGECPPPYTVYAYANSLQRTVATAQFFITGAFPGC  132

Qy    131  AITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGHRQTAFRELER  190
               |||||  ||::|||:||              :::|:||:
Db    133  DIPVHHQEKMGTMDPTFNPVITD----DSAAFSEQAVAAMEKELSKL--QLTDSYQLLEK  186

Qy    191  VLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLTEIFLLQQAQG  250
               ::|:|                    :|
Db    187  IVNYKDSPAC----KEKQQCSLVDGKNTFSAKYQQEPGVSGPLKVGNSLVDAPTLQYYEG  242

Qy    251  MP--EPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAALTPHPPQK   308
               :                         :|||||||||          |   |
Db    243  FPMDQVAWGEIKSDQQWKVLSKLKNGYQDSLFTSPEVARNVAKPLVSYIDKAL-------  295

Qy    309  QAYGVTLPTS---VLFIAGHDTNLANLGGALELN-WTLPGQPDNTPPGGELVFERWRRLS  364
               ::||||          ::||||||||||||  :||||||||||:|||||||||
Db    296  ----VTDRTSAPKITVLVGHDSNIASLLTALDFKPYQLHDQNERTPIGGKIVFQRWHDSK  351

Qy    365  DNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQIVNE  424
               ::|:    ::||:||||||||||:||||||||||||||||||||||||||||
Db    352  ANRDLMKIEYVYQSAEQLRNADALTLQAPAQRVTLELSSGC-PIDADGFCPMDKFDSVLNE  410

Qy    425  A  425
               |
Db    411  A  411


RESULT 9
E87316
periplasmic phosphoanhydride phosphohydrolase [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: E87316
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
  B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: E87316
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-414 <STO>
A;Cross-references: UNIPROT:Q9AAQ4; UNIPARC:UPI00000C70A0; GB:AE005673; NID:g13421731; P
C;Genetics:
A;Gene: CC0542

Query Match                    22.8%;    Score 526;  DB 2;  Length 414;
Best Local Similarity   34.7%;    Pred. No. 2.1e-34;
Matches  148;  Conservative   57;  Mismatches  193;  Indels   28;  Gaps    12;

Qy      9  LSL---LIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWPTWPVK   65
                     :|||||| :::::|||||:: ||||::|
Db      3  LSLRIAVVTLALASAGAASAE--TLEKVVILSRHGVRSAMSSPERLEEASARPWPRFEVP   60

Qy     66  LGHLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGEAFAAG  125
                |:|||||||||| ||||||||||||||||||||
Db     61  AGHLTARGETLVARMGDYYRRHYAAQGLLKPGDC--ASVYAWANVTQRTIATAKAYRET  117

Qy    126  LAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGHRQTAF  185
                    ||:|:|||::|||:||||||||||||||||||||||||||
Db    118  LAPGCPVTVNTVGB-GNIDPMFEPVKAGIVKADHALARAAAVAGRVGDLTAWSASHNQEA  176
```

(RIGHT COLUMN)

```
Qy    186  RELERVLNFPQSNLCL-----KREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLTE  241
               :|:|:                                           ||  |:|||
Db    177  EQLDALLMQCKGPCPPAPGKRRVFDAKPGFVDG--EEL------AGLSGPEAFASGVTE   228

Qy    242  IFLLQQAQGMPEPGWG-RITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300
               :|:|
Db    229  SLLMAWADGRDFAGLGWKSLDEEALTRSFLHQAEFDLRLRTPYVARTLAGHLADRLAAT  288

Qy    301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPG-QPDNTPPGGELVFER  359
               |                                                     |||
Db    289  LR---DGAAAIG-PVDARLVIIAGHDGTLASLGGLLRMEWTLPGYQPNQIQPGGALVFER  344

Qy    360  WRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFT  419
               :|:                  :::|:|:
Db    345  WRR-DDGVRVRVRFTGQSLSQLRNMTALDAKTPPLSAPVFVQGCGTATPAFPDCRLEDFE  403

Qy    420  QIVNEA  425
               :
Db    404  TVVRGA  409


RESULT 10
AG0632
glucose-1-phosphatase precursor (G1Pase), secreted [imported] - Salmonella enterica subs
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AG0632
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AG0632
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-413 <PAR>
A;Cross-references: UNIPARC:UPI0000059FE6; GB:AL513382; PIDN:CAD08242.1; PID:g16502289;
C;Genetics:
A;Gene: STY1153

Query Match                    22.8%;    Score 524;  DB 2;  Length 413;
Best Local Similarity   33.0%;    Pred. No. 3e-34;
Matches  135;  Conservative   65;  Mismatches  185;  Indels   24;  Gaps    10;

Qy     22  AQSEPE-LKLESVVIVSRHGVRAP-TKATQLMQDVTPDAWPTWPVKLGWLTPRGGELIAY   79
               ::        ::|:||||||||      ::||||||||||||||||||||||
Db     22  AQTTPEGYQLQQVLMMSRHNLRAPLANNGNVLAQSTPNAWPAWDVPGGQLTTKKGGVLEVV   81

Qy     80  LGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGEAFAAGLAPDCAITVHTQAD  139
               |||| |::             ::|                       :|:
Db     82  MGHYTREWLVAQGLIPSGECPAPDTVYAYANSLQRTVATAQFFITSAFPGCDIPVHHQEK  141

Qy    140  TSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNL  199
               |::|||:|| |     ::|::|::              :::|:|:
Db    142  MGTMDPTFNPVITDDSAAFRQQAVQA-MEKARSQL-----HLDESYKLLEQITHYQDSPS  195

Qy    200  CLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMP--EPGWG  257
               |                                                     :|
Db    196  C----KEKHQCSLIDAKDTFSANYQQEPGVQGPLKVGNSLVDAFTLQYYEGFPMDQVAWG  251

Qy    258  RITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAALTPHPPQKQAYGVTLPT  317
               |                          :|||||          ||
Db    252  GIHTDRQWKVLSKLKNGYQDSLFTSPTVARNVAAPLVKYIDKVLV------AERVSAP-  303

Qy    318  SVLFIAGHDTNLANLGGALELN-WTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVF  376
               ::||||||||:||:||||                :|||||||:|
Db    304  KVTVLVGHDSNIASLLTALDFKPYQLHDQYERTPIGGQLVFQRWHDGNANRDLMKIEYVY  363

Qy    377  QTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEA  425
               ::|:|:|                     :|:||:
Db    377  QTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEA  425
```

Db　　　364　QSARQLRNAEALTLKSPAQRVTLELKGCPV-DANGFCPLDKFPDNVMNTA 411

RESULT 11
S06167
acid phosphatase (EC 3.1.3.2) ACP2 precursor [validated] - human
N;Alternate names: acid phosphatase, lysosomal
C;Species: Homo sapiens (man)
C;Date: 04-Dec-1992 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004
C;Accession: S06167; S05525; S01155
R;von Figura, K.
submitted to the EMBL Data Library, June 1989
A;Reference number: S06167
A;Accession: S06167
A;Molecule type: DNA
A;Residues: 1-423 <VON>
A;Cross-references: UNIPROT:P11117; UNIPARC:UPI0000131FB1; EMBL:X15525; NID:g34239; PIDN
R;Geier, C.; von Figura, K.; Pohlmann, R.
Eur. J. Biochem. 183, 611-616, 1989
A;Title: Structure of the human lysosomal acid phosphatase gene.
A;Reference number: S05525; MUID:89377828; PMID:2776754
A;Accession: S05525
A;Molecule type: DNA
A;Residues: 1-29 <GBI>
A;Cross-references: UNIPARC:UPI00001727CD
R;Pohlmann, R.; Krentler, C.; Schmidt, B.; Schroeder, W.; Lorkowski, G.; Culley, J.; Mer
EMBO J. 7, 2343-2350, 1988
A;Title: Human lysosomal acid phosphatase: cloning, expression and chromosomal assignmen
A;Reference number: S01155; MUID:89052645; PMID:3191910
A;Accession: S01155
A;Molecule type: mRNA
A;Residues: 1-423 <POH>
A;Cross-references: UNIPARC:UPI0000131FB1; EMBL:X12548; NID:g34262; PIDN:CAA31064.1; PID
A;Note: part of this sequence, including the amino end of the mature protein, was confir
C;Genetics:
A;Gene: GDB:ACP2
A;Cross-reference: GDB:118963; OMIM:171650
A;Map position: 11p11.2-11p11.11
A;Introns: 38/3; 70/3; 99/3; 150/3; 183/3; 213/3; 258/1; 285/3; 321/2; 380/1
C;Function:
A;Description: catalyzes the hydrolysis of a wide range of phosphate esters
C;Superfamily: mammalian acid phosphatase
C;Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydrola
F;1-30/Domain: signal sequence #status predicted <SIG>
F;31-423/Product: acid phosphatase ACP2 #status experimental <MAT>
F;41/Active site: Arg #status predicted
F;42/Active site: His (phosphohistidine intermediate) #status predicted
F;92,133,167,191,267,322,331/Binding site: carbohydrate (Asn) (covalent) #status pre
F;159-370,212-310,345-349/Disulfide bonds: #status predicted

Query Match　　　6.1%;　　Score 139.5;　DB 1;　Length 423;
Best Local Similarity　23.1%;　　Pred. No. 0.0023;
Matches　98;　　Conservative　44;　Mismatches　180;　Indels 103;　Gaps 19;

Qy　　30　LESVVIVSRHGVRAPTKATQLMQDVTPDAWPTWPVKLGWLTPRGGELIAYLGHYQRQRLV 89
　　　　　　　:|::|||：|||||｜｜｜｜:||||||
Db　　33　LRFVTLLYRHGDRSPVKT---YPKDPYQEEEWPQGFGQLTKEGMLQHWELGQALRQRY- 87

Qy　　90　ADGLLAKKGCPQSGQVAIIADVD-ERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPL-- 146
　　　　　　　||：　　：||
Db　　88　-HGFL---NTSYHRQEVYVRSTDFDRTLMSAEANLAGLFP------------PNGMQR 129

Qy　　147　FNP-LKTGVCQLDNANVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREK 205
　　　　　　　||：|：||
Db　　130　FNPNISWQPIFVHTVPITE------------DRLLKFPLGPCPRYEQL 165

Qy　　206　QDESCSLTQALPSELKVSADNVSLTGAVSLASMLTEIFL----------LQQAQGMPEP 254
　　　　　　　||：||　||：||||　：|：||
Db　　166　QNE----TRQTPEYQNESSRNAQFLDMVANETGLTDLITLETVWNVYDTLFCEQTHGLRLP 221

Qy　　255　GWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAALTPHPPQKQAYGVT 314
　　　　　　　：|||　||：||||：||　：|||：|
Db　　222　PWASPQTMQRLSRLKDFSFRFLFGIYQQAEKARLQGGVLLAQIRKNLTLMATTSQ----- 276

RESULT 13
A33395

315　LPTSVLFIAGHDTNLANLGGALELNWTLPGGQPDNTPPGGELVFERWRRLSDNSQWIQVSL 374
　　　　||||：||　　｜｜｜｜｜
277　LP-KLLVYSAHDTTLVALQMALDVY-----NGEQAPYASCHIPELYQEDSGN---FSVEM 327

375　VFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIP-----A 429
　　　　：||||：|｜｜｜｜：|：　|：|
328　YF---RNESDKAPWPLSLP--------GCPHR-----CPLQDFLRL-TEPVVPKDWQQE 369

430　CSLRS 434
　　　　｜｜｜
370　CQIAS 374

RESULT 12
T16058
hypothetical protein F13D11.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change
C;Accession: T16058
R;Fulton, L.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid F13D11.
A;Reference number: S69020
A;Accession: T16058
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-416 <FUL>
A;Cross-references: UNIPROT:Q19390; UNIPARC:UPI000007AD42; EMBL:U40939; NID:g1072175; PI
C;Genetics:
A;Gene: CESP:F13D11.1
A;Introns: 23/3; 59/3; 88/3; 138/3; 202/3; 244/1; 270/3; 312/3; 349/3; 377/1
C;Superfamily: mammalian acid phosphatase

Query Match　　　5.8%;　　Score 133;　DB 2;　Length 416;
Best Local Similarity　23.4%;　　Pred. No. 0.0076;
Matches　104;　　Conservative　60;　Mismatches 174;　Indels 106;　Gaps 24;

Qy　　1　MKAILIPFLSLLIPLITPQSAPAFAQSEPELKLESSVVIVSRHGVRAPTKATQLMQDVTPD--- 57
　　　　　　：：|：|：　｜｜：|
Db　　1　MRLVLLLF--FLFPV---AFG-----KLKFVQIWFRHGERTPGHYLYFPGDDLNNVDY 48

Qy　　58　---AWPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDER 114
　　　　　　｜｜：｜　：|||
Db　　49　QQIAWP---------GELTKRGILEEFQLG--QRLLRKIYGEHFGDTYQPRDFHV--YTGKDNR 98

Qy　　115　TRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSI 174
　　　　　　：：||
Db　　99　TSASAQMFAGFLPP-------NEDQTWNYELKWQP---VAQLTDESIDWVSL----GAI 144

Qy　　175　--ADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGA 232
　　　　　　||：　：|
Db　　145　DNCPVVGEAQRKSSEYAEVMD---------QMEKYD--AELLQLVRNH----ADE-PIVEA 189

Qy　　233　VSLASMLTEI--FLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSR 289
　　　　　　：：　：||
Db　　190　VKYNHVIDSLKVRYILQDDR-LPYPEWARGYENRILNMSFLIHDA------VVKVQ 238

Qy　　290　ATPLLD---LIMAALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQ 345
　　　　　　：||
Db　　239　NDSVGDYHNELVMSYFETHLQKNSTKGV-------FISGHDTNLVTIWESLRID---GH 287

Qy　　346　PDNTPP-GGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGC 404
　　　　　　|：|　　　｜｜｜：|
Db　　288　PEDIPNYGAHIAIE---------MIEPVGQLSIKFFLSMGFNQTRVELHPHFC 331

Qy　　405　EERNAQGMCSLAGFTQIVNEARIP 428
　　　　　　｜｜：|　：|:|
Db　　332　-SRSQNNDCTWDEFQRLVKKSRKP 354

acid phosphatase (EC 3.1.3.2) precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 29-Jan-1990 #sequence_revision 29-Jan-1990 #text_change 09-Jul-2004
C;Accession: A33395
R;Himeno, M.; Fujita, H.; Noguchi, Y.; Kono, A.; Kato, K.
Biochem. Biophys. Res. Commun. 162, 1044-1053, 1989
A;Title: Isolation and sequencing of a cDNA clone encoding acid phosphatase in rat liver
A;Reference number: A33395; MUID:89350910; PMID:2764916
A;Accession: A33395
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-423 <HIM>
A;Cross-references: UNIPROT:P20611; UNIPARC:UPI0000131FE2; GB:M27893; NID:g202933; PIDN:
C;Superfamily: mammalian acid phosphatase
C;Keywords: phosphohistidine; phosphoprotein; phosphoric monoester hydrolase
F;41/Active site: Arg #status predicted
F;42/Active site: His (phosphohistidine intermediate) #status predicted

Query Match   5.7%; Score 131; DB 2; Length 423;
Best Local Similarity   22.8%; Pred. No. 0.011;
Matches 105; Conservative 49; Mismatches 184; Indels 122; Gaps 23;

Qy   2 KAILIPFL--SLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAW 59
            :||||| : | | |||||||||||||||||||||||||||||||:||||||||
Db   10 QAALLQFLLGMCLMVMPPIQA------RSLRFVTLLYRHGDRSPVKA----YPKDPYQE 58

Qy   60 PTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVD-ERTRKT 118
            ||||||||||||||||||||||||||||||||||||||||||||||||||||| :|||||
Db   59 EKWPQGFGQLTKEGMLQHWELGQALRQRY--HGFL---NASYHRQEVYVRSTDFDRTLMS 113

Qy   119 GEAFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFT 178
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   114 AEANLAGLFPP-----TEVQHFNPNISWQPI----PVHTVPITE------------ 148

Qy   179 GHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLITGAVSLASM 238
            :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   149 ---------DRLLKFPLGPCPRYEQLQNE----TRQTPEYQNMSIQNAQFLDMVANETG 194

Qy   239 LTEIFL-------LQQAQGMPEPGWGRITDSHQWNTLLSLHHNAQF-YLLQRTPEVA 286
            |||||| ||||||||||||||||||||||||||||||:||:||||||
Db   195 LMNLTLETIWNVYDTLFCEQTHGLLLPPW----ASPQTVQALSQLKDFSFLFLGIHDQVQ 251

Qy   287 RSR------ATPLLDLIMAALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELN 339
            ||| ||:|:|||||||||||||||||||||||||||||||||||||||||||
Db   252 KARLQGGVLLAQILKNLTLMATTSQFPK----------LLVYSAHDTTIVALQMALNVY 300

Qy   340 WTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKL 399
            ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ----NGKQAPYASCHIFBLYQ--EDNGNF-SVEMYFRNDSK---KAPW----------PL 340

Qy   400 TLAGCEERNAQGMCSLAGFTQIVNEARIP-----ACSLRS 434
            |||| |||||||||||||||||||||
Db   341 TLPGCPHR-----CPLQDFLRL-TEPVIPKDWQKECQLAS 374

RESULT 14
S14742
acid phosphatase (EC 3.1.3.2) precursor - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S14742
R;Geier, C.; von Figura, K.; Pohlmann, R.
Biol. Chem. Hoppe-Seyler 372, 301-304, 1991
A;Title: Molecular cloning of the mouse lysosomal acid phosphatase.
A;Reference number: S14742; MUID:91282986; PMID:2059337
A;Accession: S14742
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-421 <GEI>
A;Cross-references: UNIPROT:P24638; UNIPARC:UPI0000027103; EMBL:X57199; NID:g52870; PIDN
C;Superfamily: mammalian acid phosphatase
C;Keywords: lysosome; phosphoric monoester hydrolase

Query Match   5.3%; Score 122.5; DB 2; Length 421;
Best Local Similarity   21.8%; Pred. No. 0.054;
Matches   97; Conservative 49; Mismatches 191; Indels 107; Gaps 19;

Qy   2 KAILIPF-----LSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTP 56
            :|||| :| | |||||||||||||||||||||||||||||||:
Db   8 QAALLQFLLGMCLTVMPPIQARS--------LRFVTLLYRHGDRSPVKT----YPKDP 53

Qy   57 DAWPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVD-ERT 115
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   54 YQBEKWPQGFGQLTKEGMLQHWELGQALRQRY---NTSYHRQEVYVRSTDFDRT 108

Qy   116 RKTGEAFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIA 175
            |||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db   109 LMSAEANLAGLFPP------NEVQHFSPNISWQPI-----PVHTVPITE- 146

Qy   176 DFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLITGAVSL 235
            :||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   147 -----------DRLLKFPLGPCPRYEQLQNE----TRQTPEYQNRSIQNAQFLNMVAN 189

Qy   236 ASMLTEIFL----------LQQAQGMPEPGWGRITDSHQWNTLLSLHHNAQFYLLQRTPE 284
            |||||| ||||||||||||||||||||||||||||||:||||||||
Db   190 ETGLTNVTLETIWNVYDTLFCEQTHGLLLPPWASPQTVQRLSQLKDFSFLFLGIHEQVQ 249

Qy   285 VARSRATPLLDLIMAALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPG 344
            ||||||||||||||||||||||||||||||||||||||||||||||
Db   250 KARLQGGVLLAQILKNLTLMATTSQF--------PKLLVYSAHDTTIVALQMALNVY-- 298

Qy   345 QPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGC 404
            |||||||||||||||||||||||||||||||||||||||||||||||
Db   299 NGKQAPYASCHIFBLYQ--EDNGNF-SVEMYFRNDSK--KAPWPLILP---------GC 343

Qy   405 EERNAQGMCSLAGFTQIVNEARIP 428
            |||||||||||||||||||||
Db   344 PHR------CPLQDFLRL-TEPVIP 361

RESULT 15
JH0610
acid phosphatase (EC 3.1.3.2) ACPP precursor [validated] - human
N;Alternate names: acid phosphatase, prostatic; orthophosphoric monoester phosphohydrola
C;Species: Homo sapiens (man)
C;Date: 17-Aug-1992 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004
C;Accession: JH0610; JS0693; A38608; S01331; A32419; S11147; S38863; S41251; S17042; S42
R;Sharief, F.S.; Li, S.S.L.
Biochem. Biophys. Res. Commun. 184, 1468-1476, 1992
A;Title: Structure of human prostatic acid phosphatase gene.
A;Reference number: JH0610; MUID:92272747; PMID:1375464
A;Accession: JH0610
A;Molecule type: DNA
A;Residues: 1-386 <SHA>
A;Cross-references: UNIPARC:UPI000000000B; UNIPROT:P15309; UNIPARC:UPI0000039832; GB:M97580; GB:M975
A;Accession: JS0693
A;Molecule type: mRNA
A;Residues: 1-386 <SH3>
A;Cross-references: UNIPARC:UPI0000039832; GB:M97589; NID:g189611; PIDN:AAA60021.1; PID:
R;Van Etten, R.L.; Davidson, R.; Stevis, P.E.; MacArthur, H.; Moore, D.L.
J. Biol. Chem. 266, 2313-2319, 1991
A;Title: Covalent structure, disulfide bonding, and identification of reactive surface a
A;Reference number: A38608; MUID:91115848; PMID:1989985
A;Accession: A38608
A;Molecule type: mRNA
A;Residues: 1-386 <VAN>
A;Cross-references: UNIPARC:UPI0000039832; GB:M34840; NID:g189620; PIDN:AAA69694.1; PID:
A;Note: part of this sequence, including the amino end of the mature protein, was confir
R;Vihko, P.; Virkkunen, P.; Henttu, P.; Roiko, K.; Solin, T.; Huhtala, M.L.
FEBS Lett. 236, 275-281, 1988
A;Title: Molecular cloning and sequence analysis of cDNA encoding human prostatic acid p
A;Reference number: S01331; MUID:88312981; PMID:2842184
A;Accession: S01331
A;Molecule type: mRNA
A;Residues: 1-14,'A',16,'ASC',20,'CF',23,'C',25-65,'WIWPTHPA',74-211,'A',213-386,<VIH>
A;Cross-references: UNIPARC:UPI00016A48C; EMBL:X52174; NID:g28321; PIDN:CAA36422.1; PID

F;1-32/Domain: signal sequence #status predicted <SIG>
F;33-386/Product: acid phosphatase ACPP #status experimental <MAT>
F;43,86/Active site: Arg #status predicted
F;44/Active site: His (phosphohistidine intermediate) #status predicted
F;94,220,333/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;161-372,215-313,347-351/Disulfide bonds: #status experimental

Query Match   5.1%;   Score 117;   DB 1;   Length 386;
Best Local Similarity   19.8%;   Pred. No. 0.13;
Matches   87;   Conservative   63;   Mismatches 178;   Indels 112;   Gaps   15;

Qy   3 AILIPFLSLLIPLTPQSAFAQSBPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWT-   61
              :: | :|| |||| || |:: | || | | | | ||
Db   13 SLSLGFLFLLFFWLDRSVLAK----ELKFVTLVFRHGDRSPI---------DTFPTD   56

Qy   62 -----WPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQS--GQVAIIADVD 112
                 ||| |||| | ||| |:| |: | |||
Db   57 PIKESSWPQGFGQLTQLGMEQHYELGSYIRKRY------RKFLNESYKHEQYIRSTDV 109

Qy   113 ERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGG 172
           :|| :   :   : |
Db   110 DRTLMSAMTNLAALFPPBGVSIW------NPILLWQPIPVHTVPLSE---------- 150

Qy   173 SIADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGA 232
                                 --DQLLYLPFRN-CPRFQELESETLKSBEFQKRLHPYKDFIATLGK 193
Db   151 ---------------------

Qy   233 VS--------LASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTP 283
            :| : | |: | : ::   :   : :.
Db   194 LSGLHGQDLFGIWSKVYDPLYCESVHNFTLFSWATEDTMTKLRELSELSLLSLYGIHKQK 253

Qy   284 EVARSRATPLLDLIMAALTPHPPQKQAYGVTLPTSVLFIAGHDTNIANLGGALEL-NWTL 342
            | ::   : || |:| | :: || : || |:: : :|
Db   254 EKSRLQGGVLVNEILNHM------KRATQIPSYKKLIMYSAHDTTVSGLQMALDVYNGLL 307

Qy   343 PCQPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLA 402
           | | --VEMYYRN---------ETQHEPYPLMLP 345
Db   308 P--PYASCHLTBLYFBKGBYF------

Qy   403 GCEBRNAQGMCSLAGFTQIV 422
           || |||| ||:|
Db   346 GCSP-----SCPLERFAELV 360

Search completed: June 13, 2006, 10:29:29
Job time : 28.7339 secs

A;Note: part of this sequence, including the amino end of the mature protein, was confi
R;Sharief, F.S.; Lee, H.; Leiderman, M.M.; Lundwall, A.; Deaven, L.L.; Lee, C.; Li, S.S.
Biochem. Biophys. Res. Commun. 160, 79-86, 1989
A;Title: Human prostatic acid phosphatase: cDNA cloning, gene mapping and protein sequen
A;Reference number: A32419; MUID:89228054; PMID:2712834
A;Accession: A32419
A;Molecule type: mRNA
A;Residues: 1-14,'A',16,'ASC',20,'CF',23,'C',25-94,'D',96-115,'R',117-214,'S',216-293,'T
A;Cross-references: UNIPARC:UPI00016AB62; GB:M24902; NID:g189618; PIDN:AAA60022.1; PID:
A;Note:the authors translated the codons GAC for residue 95 as Glu, CGT for residue 116
R;Tailor, P.G.; Govindan, M.V.; Patel, P.C.
Nucleic Acids Res. 18, 4928, 1990
A;Title: Nucleotide sequence of human prostatic acid phosphatase determined from a full-
A;Reference number: S11147; MUID:90370491; PMID:2395659
A;Accession: S11147
A;Status: preliminary; translation not shown
A;Molecule type: mRNA
A;Residues: 1-14,'AFASC',20,'CF',23,'C',25-45,'H',47-65,'RIWPTHPA',74-138,'E',140-156,'R
A;Cross-references: UNIPARC:UPI00016AF1A; EMBL:X53605; NID:g35683; PIDN:CAA37673.1; PID
R;Banas, B.; Blaschke, D.; Fittler, F.; Hoerz, W.
submitted to the EMBL Data Library, April 1993
A;Description: Characterization of the promoter of the human prostatic acid phosphatase
A;Reference number: S38863
A;Accession: S38863
A;Molecule type: DNA
A;Residues: 1-40 <BAN>
A;Cross-references: UNIPARC:UPI0000006BB; EMBL:X71391; NID:g416530; PIDN:CAA50514.1; PI
R;Virkkunen, P.H.; Hedberg, P.H.; Palvimo, J.J.; Birr, E.; Porvari, K.; Taavitsainen, P.;
submitted to the EMBL Data Library, September 1993
A;Description: Structural organization of human and rat prostate-specific acid phosphata
equence in the human gene promoter.
A;Reference number: S41251
A;Accession: S41251
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-40 <VIR>
A;Cross-references: UNIPARC:UPI0000006BB; EMBL:X74961; NID:g439657; PIDN:CAA52913.1; PI
R;Lee, T.M.; Li, S.S.L.; Lee, C.
Biochem. J. 277, 759-765, 1991
A;Title: Homodimer and heterodimer subunits of human prostate acid phosphatase.
A;Reference number: S17042; MUID:91336999; PMID:1908222
A;Accession: S17042
A;Status: preliminary
A;Molecule type: protein
A;Residues: 33-49 <LEE>
A;Cross-references: UNIPARC:UPI00001727CC
R;Banas, B.; Blaschke, D.; Fittler, F.; Hoerz, W.
Biochim. Biophys. Acta 1217, 188-194, 1994
A;Title: Analysis of the promoter of the human prostatic acid phosphatase ACPP gene.
A;Reference number: I37175; MUID:95038536; PMID:7951074
A;Accession: I37175
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-386 <RES>
A;Cross-references: UNIPARC:UPI0000039832; EMBL:U07097; NID:g515995; PIDN:AAB60640.1; PI
C;Comment: This protein is synthesized under androgen regulation by epithelial cells of
C;Genetics:
A;Gene: GDB:ACPP
A;Cross-references: GDB:119644; OMIM:171790
A;Map position: 3q21.3-3q25.2
A;Introns: 40/3; 72/3; 101/3; 152/3; 185/3; 216/3; 260/3; 288/3; 323/2
C;Function:
A;Description: catalyzes the hydrolysis of a wide range of phosphate esters
C;Superfamily: mammalian acid phosphatase
C;Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydrola

```
                    GenCore version 5.1.9
              Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on:        June 13, 2006, 10:46:00 ; Search time 10.0917 Seconds
                                (without alignments)
                              551.244 Million cell updates/sec

Title:          US-10-601-319-2
Perfect score:  2302
Sequence:       1 MKAILIPFLSLLIPLTPQSA..........IVNEARIPACSLRSHHHHHH 440

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       64916 seqs, 12643201 residues

Total number of hits satisfying chosen parameters:   64916

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      Published Applications AA New:*
                1: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
                2: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
                3: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
                4: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
                5: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
                6: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
                7: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
                8: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

                              SUMMARIES

Result No.	Score	Query Match %	Query Match Length	DB	ID	Description
1	98	4.3	1730	7	US-11-270-040-6	Sequence 6, Appli
2	97	4.2	460	6	US-10-953-349-34904	Sequence 34904, A
3	97	4.2	464	6	US-10-953-349-34903	Sequence 34903, A
4	97	4.2	472	6	US-10-953-349-34902	Sequence 34902, A
5	93.5	4.1	4391	7	US-11-183-325-56	Sequence 56, Appli
6	91.5	4.0	635	6	US-10-511-937-2424	Sequence 2424, Ap
7	91	4.0	458	6	US-10-953-349-8297	Sequence 8297, Ap
8	91	4.0	613	6	US-10-953-349-8296	Sequence 8296, Ap
9	91	4.0	614	6	US-10-953-349-8295	Sequence 8295, Ap
10	90	3.9	384	7	US-11-293-697-2758	Sequence 2758, Ap
11	87.5	3.8	952	7	US-11-293-697-3284	Sequence 3284, Ap
12	86	3.7	1336	6	US-10-511-937-2564	Sequence 2564, Ap
13	84	3.6	370	6	US-10-196-749-350	Sequence 350, App
14	84	3.6	725	7	US-11-293-697-3258	Sequence 3258, Ap
15	82.5	3.6	360	6	US-10-953-349-22123	Sequence 22123, A
16	81.5	3.5	838	7	US-11-295-336-2	Sequence 2, Appli
17	81	3.5	281	6	US-10-953-349-23762	Sequence 23762, A
18	80.5	3.5	5738	6	US-10-505-928-150	Sequence 150, App
19	79.5	3.5	269	6	US-10-953-349-7	Sequence 7, Appli
20	79.5	3.5	282	6	US-10-953-349-6	Sequence 6, Appli
21	79	3.4	505	7	US-11-293-697-3134	Sequence 3134, Ap
22	79	3.4	826	7	US-11-301-554-330	Sequence 330, App
23	79	3.4	1193	6	US-10-505-928-537	Sequence 537, Appl
24	78.5	3.4	765	7	US-11-317-329-3	Sequence 3, Appli
25	78.5	3.4	765	7	US-11-317-329-34	Sequence 34, Appl
26	78.5	3.4	765	7	US-11-317-329-35	Sequence 35, Appl
27	78.5	3.4	765	7	US-11-317-329-36	Sequence 36, Appl
28	78.5	3.4	765	7	US-11-317-329-37	Sequence 37, Appl
29	78.5	3.4	765	7	US-11-317-329-38	Sequence 38, Appl
30	78.5	3.4	765	7	US-11-317-329-39	Sequence 39, Appl
31	78.5	3.4	765	7	US-11-317-329-40	Sequence 40, Appl
32	78.5	3.4	765	7	US-11-317-329-41	Sequence 41, Appl
33	78.5	3.4	765	7	US-11-317-329-42	Sequence 42, Appl
34	78.5	3.4	765	7	US-11-317-329-43	Sequence 43, Appl
35	78.5	3.4	765	7	US-11-317-329-44	Sequence 44, Appl
36	78.5	3.4	1043	7	US-11-299-697-3024	Sequence 3024, Ap
37	78	3.4	9535	6	US-10-471-571A-4496	Sequence 4496, Ap
38	77.5	3.4	307	6	US-10-953-349-38096	Sequence 38096, A
39	77.5	3.4	351	6	US-10-953-349-22620	Sequence 22620, A
40	77	3.3	341	6	US-10-953-349-22124	Sequence 22124, A
41	77	3.3	384	6	US-10-953-349-1524	Sequence 1524, Ap
42	77	3.3	460	6	US-10-953-349-1523	Sequence 1523, Ap
43	77	3.3	485	6	US-10-953-349-32471	Sequence 32471, A
44	77	3.3	502	6	US-10-953-349-1522	Sequence 1522, Ap
45	77	3.3	504	6	US-10-953-349-32470	Sequence 32470, A

                              ALIGNMENTS

```
RESULT 1
US-11-270-040-6
; Sequence 6, Application US/11270040
; Publication No. US20060110762A1
; GENERAL INFORMATION:
;  APPLICANT: Kapil, Sanjay
;  APPLICANT: Kim, Jeong-Ki
;  TITLE OF INVENTION: PORCINE REPRODUCTIVE AND RESPIRATORY
;  TITLE OF INVENTION: SYNDROME VIRUS RECEPTOR COMPONENTS AND USES THEREOF
;  FILE REFERENCE: 14337/004001
;  CURRENT APPLICATION NUMBER: US/11/270,040
;  CURRENT FILING DATE: 2005-11-09
;  PRIOR APPLICATION NUMBER: 60/680,297
;  PRIOR FILING DATE: 2005-05-12
;  PRIOR APPLICATION NUMBER: 60/626,788
;  PRIOR FILING DATE: 2004-11-10
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;  LENGTH: 1730
;  TYPE: PRT
;  ORGANISM: Sus scrofa
US-11-270-040-6

Query Match        4.3%;  Score 98;  DB 7;  Length 1730;
Best Local Similarity   21.2%;  Pred. No. 2.1;
Matches   94;  Conservative   54;  Mismatches  184;  Indels  112;  Gaps   20;

Qy     22 AQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWPTWPVKLGWLTPRGGELIAYLG 81
          |:||||:||:|||||||:||||||||||:||||||||||||||||||||—||||||||||
Db   1250 ANTSLELRLEGVQVALAPSATVPEGAPVTVTCEDPAARP--PTLYVW            1294

Qy     82 HYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGEAFAAGLAPDCAITV-HTQAD- 139
          :|:::||:::::|::|::||:||::::::::|::::::::|:||::||:::|::::::::|
Db   1295 -YHNSRWLQEGSAASLSFP----AATRAHAGAYTCQVQDAQGTRISQPAALHILYAPRDA 1349

Qy    140 --TSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADF---------TGHRQTAFREL  188
            :|::::::::::::::::|::::|::|||:::::||::::::      :||:|:||:|
Db   1350 VLSSFWDSRASPMAVVQCTVDSEPPAEMTLSHDGKVLATSHGVHGLAVGTGHVQVARNAL 1409

Qy    189 E-RVLNFPQSN---LCLKREKQDESCSLTQALPSELKVSAD---NVSLTGAVSLASMLT  240
          ::::|::::::::::::::::::|:|::::::::|:::::|::::::::::|:::|::::
Db   1410 QLRVQNVPSRDKDTYVCMDRNSLGSVSTMGLQPEGVHVVAEPGLDVPEGTALNLSCRLP 1469

Qy    241 EIFLLQQAQGMPEPG---------W---GRITDSHQWNTLLSLHNAQ-----FYLLQRTPE 284
          ::::::::::|::::          :   ||::::::::::::::|      :|::||:||
Db   1470 S---------GPGHIGNSTFAWFRNGRQLHTESVPTLTFTHVARAQAGLYHCQAELPA    1518
```

```
Qy    285  VARSRATPLLDLIMAALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPG  344
                ::|   |||:::    |:||   ::|:|
Db   1519  GAATSAPVLLRVL------YPPK-------TPTMTVFVEPE---GG--IQGILDC      1555

Qy    345  QPDNTP------PGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEV   397
                ||||                    ::|:|||
Db   1556  RVDSBFLASLTLHLGSRLVASSQPQAAPAKPHIRVS----------ASPNALRVDMEEL  1604

Qy    398  KLTLAG---CBERNAQGMCSLAGF  418
                |:|
Db   1605  KPSDQGEYVCSASNALGSASAATY 1628
```

RESULT 2
```
US-10-953-349-34904
; Sequence 34904, Application US/10953349
; Publication No. US2006010734SA1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCONDED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34904
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34904

Query Match        4.2%;    Score 97;   DB 6;   Length 460;
Best Local Similarity   25.1%;   Pred. No. 0.38;
Matches 81;   Conservative 37;   Mismatches 107;   Indels 98;   Gaps 23;
```

```
Qy   170  AGGSIADFT---------GHRQTAF--RELERVLNFPQS---------NLCLKR-EKQDE  208
               ||||||           :|:          |::              :|
Db     6  AFGHILPFTELAKRIARQGHRVTLFSTPRNTRRLIRIPPELAGHIRVVDIALPRVERLPE   65
Qy   209  SCSLTQALPSE-----LKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPE-PGWGRITDS  262
               |:|:|         :|::                     |:|:|
Db    66  DCEASIDLPSDDLRPYLRVAYD------AAFADKLSAI--LQEPR--PERPDWVLIDYA   114
Qy   263  HQWNTLLSLHNAQFYLLQRTPEVARSRATP--LLDLIMAALTPH--PPQ---KQAYGVT   314
Db   115  AYW----------APAAAARHGVPCAFLSLFGAATLSFYGPPEGLMGRGKYART         158
Qy   315  LPTSVLFIAGH--DTNLANLG-GALELNWTLPG-QPDNT---------PPGGELVFER    359
               |||     :||     ::::
Db   159  KPEDLTVVPDYVPFPTTVAHRGFEAREL--FKPGLVPDDSGVSEGHRFGVSIGESQIVGI  216
Qy   360  WRRLSDNSQWIQVSLVFQTLQQMRDK--TPLSLNTPPGEVKLTLAGCE------ERNAQG  411
               :|                       ||::              |:|
Db   217  RSRTELESEWLQV------LGKLYKKPVIPIGLFPPP--PTQDIAGHEATLRWLDRQAQG  268
Qy   412  MCSLAGFTQIVNEARIPACSLRS  434
               :||:                ::|
Db   269  SVVYAAFG---SEAKLTSAQLQT  288
```

RESULT 3
```
US-10-953-349-34903
; Sequence 34903, Application US/10953349
; Publication No. US2006010734SA1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCONDED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34903
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34903

Query Match        4.2%;    Score 97;   DB 6;   Length 464;
Best Local Similarity   25.1%;   Pred. No. 0.38;
Matches 81;   Conservative 37;   Mismatches 107;   Indels 98;   Gaps 23;
```

```
Qy   170  AGGSIADFT---------GHRQTAF--RELERVLNFPQS---------NLCLKR-EKQDE  208
Db    10  AFGHILPFTELAKRIARQGHRVTLFSTPRNTRRLIRIPPELAGHIRVVDIALPRVERLPE   69
Qy   209  SCSLTQALPSE-----LKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPE-PGWGRITDS  262
Db    70  DCEASIDLPSDDLRPYLRVAYD------AAFADKLSAI--LQEPR--PERPDWVLIDYA   118
Qy   263  HQWNTLLSLHNAQFYLLQRTPEVARSRATP--LLDLIMAALTPH--PPQ---KQAYGVT   314
Db   119  AYW----------APAAAARHGVPCAFLSLFGAATLSFYGPPEGLMGRGKYART         162
Qy   315  LPTSVLFIAGH--DTNLANLG-GALELNWTLPG-QPDNT---------PPGGELVFER    359
Db   163  KPEDLTVVPDYVPFPTTVAHRGFEAREL--FKPGLVPDDSGVSEGHRFGVSIGESQIVGI  220
Qy   360  WRRLSDNSQWIQVSLVFQTLQQMRDK--TPLSLNTPPGEVKLTLAGCE------ERNAQG  411
Db   221  RSRTELESEWLQV------LGKLYKKPVIPIGLFPPP--PTQDIAGHEATLRWLDRQAQG  272
Qy   412  MCSLAGFTQIVNEARIPACSLRS  434
Db   273  SVVYAAFG---SEAKLTSAQLQT  292
```

RESULT 4
```
US-10-953-349-34902
; Sequence 34902, Application US/10953349
; Publication No. US2006010734SA1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCONDED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34902
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34902

Query Match        4.2%;    Score 97;   DB 6;   Length 472;
Best Local Similarity   25.1%;   Pred. No. 0.39;
Matches 81;   Conservative 37;   Mismatches 107;   Indels 98;   Gaps 23;
```

```
Qy   170  AGGSIADFT---------GHRQTAF--RELERVLNFPQS---------NLCLKR-EKQDE  208
Db    18  AFGHILPFTELAKRIARQGHRVTLFSTPRNTRRLIRIPPELAGHIRVVDIALPRVERLPE   77
Qy   209  SCSLTQALPSE-----LKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPE-PGWGRITDS  262
Db    78  DCEASIDLPSDDLRPYLRVAYD------AAFADKLSAI--LQEPR--PERPDWVLIDYA   126
Qy   263  HQWNTLLSLHNAQFYLLQRTPEVARSRATP--LLDLIMAALTPH--PPQ---KQAYGVT   314
Db   127  AYW----------APAAAARHGVPCAFLSLFGAATLSFYGPPEGLMGRGKYART         170
Qy   315  LPTSVLFIAGH--DTNLANLG-GALELNWTLPG-QPDNT---------PPGGELVFER    359
```

```
Db    171  KPEDLTVVPDYVPPPTVAHRGFEAREL--FKPGLVPDDSGVSEGHRFGVSIGESQIVGI   228

Qy    360  WRRLSDNSQWIQVSLVFQTLQQMRDK--TPLSLNTPPGPEVKLTLAGCE------ERNAQG   411
Db    229  RSRTELRSEWLQV------LGKLYKKPVIPIGLFPPP--PTQDIAGHEATLRWLDRQAQG   280

Qy    412  MCSLAGFTQIVNEARIPACSLRS   434
Db    281  SVVYAAFG---SEAKLTSAQLQT   300
```

RESULT 5
US-11-183-325-56
; Sequence 56, Application US/11183325
; Publication No. US20060104898A1
; GENERAL INFORMATION:
;   APPLICANT: Vanderbilt University
;   APPLICANT: Hallahan, Dennis E
;   APPLICANT: Qu, Shimian
;   TITLE OF INVENTION: IN VIVO PANNING FOR LIGANDS TO RADIATION-INDUCED MOLECULES
;   FILE REFERENCE: 1242/4/7/2 CIP
;   CURRENT APPLICATION NUMBER: US/11/183,325
;   CURRENT FILING DATE: 2005-07-15
;   PRIOR APPLICATION NUMBER: US 60/328123
;   PRIOR FILING DATE: 2001-10-03
;   PRIOR APPLICATION NUMBER: US 10/259,087
;   PRIOR FILING DATE: 2002-09-27
;   NUMBER OF SEQ ID NOS: 56
;   SOFTWARE: PatentIn version 3.3
; SEQ ID NO 56
;   LENGTH: 4391
;   TYPE: PRT
;   ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
;   DATABASE ACCESSION NUMBER: P98160
;   DATABASE ENTRY DATE: 2003-02-28
;   RELEVANT RESIDUES: (1)..(4391)
US-11-183-325-56

Query Match   4.1%; Score 93.5; DB 7; Length 4391;
Best Local Similarity   21.4%; Pred. No. 21; Indels 167; Gaps 25;
Matches 110; Conservative 50; Mismatches 186;

Qy   6 IPPLSLLIPLTPQSAFAQSEPELKLESV-VIVSRHGVRAPTKATQLMQDVTPDAWPTWPV   64
Db   3389 LPATSIPAGSTPT---VQVTPQLETKSIGASVEFHCAVPSDRGTQL--------RWFK   3435

Qy   65 KLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGEAFAA   124
Db   3436 EGGQLPP--------GH-----SVDDGVLRIQNLDQSCQGTYIC---QAHGPWGKAQAS   3478

Qy   125 G-----LAPDCAITVHTQADT-----------SSPDP-------LFNPLKTGVCQLDN   159
Db   3479 AQLVIQALPSVLINIRTSVQTVVVGHAVEFECLALGDPKPQVTWSKVGGGHLRPGIVQ---   3535

Qy   160 ANVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQAL----   216
Db   3536 ---SGGVRIAHVBLADAGQYRCTA----TNAAGTTQSHVLL--------LVQALPQI   3578

Qy   217 --PSELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEP--GWGRI-----TDSHQWNT   267
Db   3579 SMPQEVRVPAGSAAVPPCI--------ASGYPTDISWSKLDGSLPPDSRLENN   3624

Qy   268 LLSL------HNAQFYLLQRTPEVAR----------SRATPLLDLIMAALTPHPPQKQAY-   311
Db   3625 MLMLPSVRPQDAGTYVCTATNRQGKVKAFAHLQVPERVVPYPTQTPYSFLPTIKQDAYR   3684

Qy   312 ----------GVTLPTSVLPIAGHDTNLANLGGALELNWTLPGQPDNTPPG------   352
Db   3685 KPEIKITFRPDSADGMLLYNGQKRVPGSPTNLAN-----------RQPDPISFGLVGGR   3732

Qy   353 GELVFERWRRLSDNSQWIQVSL-VFQTLQQMRDKTPLSL----------NTPPGBV------   397
```

```
Db   3733  PEFRPDAGSGMATIRHPTPLALGHGPHTVTLLRSLTQGSLLIVGDLAPVNGTSQGKFQGLDL   3792

Qy    398  --KLTLAGCEBRNAQGMCSL-AGFTQIVNEARI   427
Db   3793  NEELYLGGYPDYGAIPAKGLSSGFPIGCVRELRI   3825
```

RESULT 6
US-10-511-937-2424
; Sequence 2424, Application US/10511937
; Publication No. US20060088836A1
; GENERAL INFORMATION:
;   APPLICANT: EXPRESSION DIAGNOSTICS, INC.
;   APPLICANT: Wohlgemuth, Jay
;   APPLICANT: Fry, Kirk
;   APPLICANT: Woodward, Robert
;   APPLICANT: Ly, Ngoc
;   APPLICANT: Prentice, James
;   APPLICANT: Morris, MacDonald
;   APPLICANT: Rosenberg, Steven
;   TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;   TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
;   FILE REFERENCE: 50661200104
;   CURRENT APPLICATION NUMBER: US/10/511,937
;   CURRENT FILING DATE: 2004-10-19
;   PRIOR APPLICATION NUMBER: PCT/US2003/012946
;   PRIOR FILING DATE: 2003-04-24
;   PRIOR APPLICATION NUMBER: US 10/131,831
;   PRIOR FILING DATE: 2002-04-24
;   PRIOR APPLICATION NUMBER: US 10/325,899
;   PRIOR FILING DATE: 2002-12-20
;   NUMBER OF SEQ ID NOS: 3117
;   SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2424
;   LENGTH: 635
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-511-937-2424

Query Match   4.0%; Score 91.5; DB 6; Length 635;
Best Local Similarity   19.8%; Pred. No. 1.9;
Matches 111; Conservative 65; Mismatches 196; Indels 189; Gaps 25;

Qy   2 KAILIPPLSLLIPLTPQSAFAQSEP-ELKL---ESVVIVS------RHGVRAPTKAT-   48
Db   123 RVLFVDSVGLPAPPSIIKAMGGSQPGELQISWEBPAPEISDFLRYELRYGPRDPKNSTGP   182

Qy   49 --QLMQ----------------------DVTPDAWPTWPVKLG-----------WLTPRGGE-L   76
Db   183 TVIQLIATETCCPALQRPHSASALDQSPCAQPTMPWQDGPKQTSPSREASALTAEGGSCL   242

Qy   77 IAYL----GHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGEAFAAGLAPDCAI   132
Db   243 ISGLQPGNSYWLQLRSEPDG-ISLGGSWGSWGSWSLPVTVDL-------PGDAVALGL----   289

Qy   133 TVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGHRQTAFRELERVL   192
Db   290 ----QCFT---------LDLKNVTCWQQQDHASSQGFFYHSRARCCPRDR--   327

Qy   193 NFPQSNLCLKREKQDESCSLTQALPSELKVSADNV-------SLTGAVSLASMLTEIFLLQ   246
Db   328 -YPIWENCEBEEKTNPGLQTPQFSRCHFKSRNDSIIHILVEVTTAPGTVHSYLGSPFWIH   386

Qy   247 QAQGMPEPG-WGRITDSH----------QWNTLLLSLHNAQF   276
Db   387 QAVRLFTPNLHWREISSGHLELEWQHPSSWAAQETCYQLRYTGEGHQDWKVLEPPLGARG   446

Qy   277 YLLQRTPEVARSRATPLLDLIMAALTPHPP----------QKQAYGVTLPTSVLPIAG   324
Db   447 GTLELRP---RSRYRLQLRARLNGPTYQGPWSSWSDPTRVETATETAWISLVTALHLVLG   503

Qy   325 HDTNLANLGGALELNWTLPGQ----------------------PD--------NTPPGGELVFE   358
```

```
Db   504  ---LSAVLGLLLRWQFPAHYRRLRHALWPSLPDLHRVLGQYLRDTAALSPPKA----- 554
Qy   359  RWRRLSDNSWQIQVSLVFQTLQQMRDKTPLSLNTPPGEV-------KLTLAGCEER 407
          : ||:|||   |  | ::|||  ||          :: :
Db   555  ---TVSDTCEEVEPSLL-EILPKSSERTPLPLCSSOAQMDYRRLQPSCLGTMPLSVCPPM 610
Qy   408  NAQGMCSLAGFTQIVNEARIP 428
          | |  :::||  |
Db   611  ABSGSCCT---THIANHSYLP 628

RESULT 7
US-10-953-349-8297
; Sequence 8297, Application US/10953349
; Publication No. US2006010734SA1
; GENERAL INFORMATION:
;    APPLICANT: ALEXANDROV, Nickolai et al.
;    TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
;    TITLE OF INVENTION:  ENCONDED THERBY
;    FILE REFERENCE: 2750-1579PUS2
;    CURRENT APPLICATION NUMBER: US/10/953,349
;    CURRENT FILING DATE:  2004-09-30
;    NUMBER OF SEQ ID NOS: 40252
;    SOFTWARE: Patentin version 3.3
; SEQ ID NO 8297
;    LENGTH: 458
;    TYPE: PRT
;    ORGANISM: Arabidopsis thaliana
US-10-953-349-8297

Query Match             4.0%;  Score 91;  DB 6;  Length 458;
Best Local Similarity   20.9%;  Pred. No. 1.3;
Matches   50;  Conservative   31;  Mismatches   92;  Indels   66;  Gaps   10;

Qy   140  TSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADF------TGHRQTAFRELERVL 192
          : ||:|||  :||  |   :
Db   216  SGSPNGLLSDIK-----FKMMNLTFRRWSNYSAGLVDYRPGCQGLVNHRATSGIIMEHVN 270
Qy   193  NFPQSNLCLKREKQDESCSLTQALPSELKVS-ADNVSLTGAVSLASMLTEIFLLQQAQGM 251
Db   271  GFRVENVDLK-WSDDDDVNAAWNVPLEFRPSTVNNVSFVGFTS--GLYTKLFESDYVMSL 327
Qy   252  PEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAALTPHPPQKQAY 311
          | |:||
Db   328  ------FTSNHQYIT------ 355
Qy   312  GVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDN-------TPPGGELVFERW 360
          ||  ||| :::: | :    RWATPGHEERPKGYFMNRTPPPGQSRKWEDW
Db   356  -KTLRNRIHFRSGSTSTGPS------RWATPGHEERPKGYFMNRTPPPGQSRKWEDW 406

RESULT 8
US-10-953-349-8296
; Sequence 8296, Application US/10953349
; Publication No. US2006010734SA1
; GENERAL INFORMATION:
;    APPLICANT: ALEXANDROV, Nickolai et al.
;    TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
;    TITLE OF INVENTION:  ENCONDED THERBY
;    FILE REFERENCE: 2750-1579PUS2
;    CURRENT APPLICATION NUMBER: US/10/953,349
;    CURRENT FILING DATE:  2004-09-30
;    NUMBER OF SEQ ID NOS: 40252
;    SOFTWARE: Patentin version 3.3
; SEQ ID NO 8296
;    LENGTH: 613
;    TYPE: PRT
;    ORGANISM: Arabidopsis thaliana
US-10-953-349-8296

Query Match             4.0%;  Score 91;  DB 6;  Length 613;
Best Local Similarity   20.9%;  Pred. No. 2;
Matches   50;  Conservative   31;  Mismatches   92;  Indels   66;  Gaps   10;
```

```
Qy   140  TSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADF------TGHRQTAFRELERVL 192
          : ||:|||  :||  |   :
Db   371  SGSPNGLLSDIK-----FKMMNLTFRRWSNYSAGLVDYRPGCQGLVNHRATSGIIMEHVN 425
Qy   193  NFPQSNLCLKREKQDESCSLTQALPSELKVS-ADNVSLTGAVSLASMLTEIFLLQQAQGM 251
Db   426  GFRVENVDLK-WSDDDDVNAAWNVPLEFRPSTVNNVSFVGFTS--GLYTKLFESDYVMSL 482
Qy   252  PEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAALTPHPPQKQAY 311
          | |:||
Db   483  ------FTSNHQYIT------ 510
Qy   312  GVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDN-------TPPGGELVFERW 360
          ||  ||| :::: | :
Db   511  -KTLRNRIHFRSGSTSTGPS------RWATPGHEERPKGYFMNRTPPPGQSRKWEDW 561

RESULT 9
US-10-953-349-8295
; Sequence 8295, Application US/10953349
; Publication No. US2006010734SA1
; GENERAL INFORMATION:
;    APPLICANT: ALEXANDROV, Nickolai et al.
;    TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
;    TITLE OF INVENTION:  ENCONDED THERBY
;    FILE REFERENCE: 2750-1579PUS2
;    CURRENT APPLICATION NUMBER: US/10/953,349
;    CURRENT FILING DATE:  2004-09-30
;    NUMBER OF SEQ ID NOS: 40252
;    SOFTWARE: Patentin version 3.3
; SEQ ID NO 8295
;    LENGTH: 614
;    TYPE: PRT
;    ORGANISM: Arabidopsis thaliana
US-10-953-349-8295

Query Match             4.0%;  Score 91;  DB 6;  Length 614;
Best Local Similarity   20.9%;  Pred. No. 2;
Matches   50;  Conservative   31;  Mismatches   92;  Indels   66;  Gaps   10;

Qy   140  TSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADF------TGHRQTAFRELERVL 192
          : ||:|||  :||  |   :
Db   372  SGSPNGLLSDIK-----FKMMNLTFRRWSNYSAGLVDYRPGCQGLVNHRATSGIIMEHVN 426
Qy   193  NFPQSNLCLKREKQDESCSLTQALPSELKVS-ADNVSLTGAVSLASMLTEIFLLQQAQGM 251
Db   427  GFRVENVDLK-WSDDDDVNAAWNVPLEFRPSTVNNVSFVGFTS--GLYTKLFESDYVMSL 483
Qy   252  PEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAALTPHPPQKQAY 311
          | |:||
Db   484  ------FTSNHQYIT------ 511
Qy   312  GVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDN-------TPPGGELVFERW 360
          ||  ||| :::: | :
Db   512  -KTLRNRIHFRSGSTSTGPS------RWATPGHEERPKGYFMNRTPPPGQSRKWEDW 562

RESULT 10
US-11-293-697-2758
; Sequence 2758, Application US/11293697
; Publication No. US2006010S376A1
; GENERAL INFORMATION:
;    APPLICANT: HELIX RESEARCH INSTITUTE
;    TITLE OF INVENTION: Novel full length cDNA
;    FILE REFERENCE: H1-A0106
;    CURRENT APPLICATION NUMBER: US/11/293,697
;    CURRENT FILING DATE:  2005-12-05
;    PRIOR APPLICATION NUMBER: US/10/108,260
;    PRIOR FILING DATE: 2002-03-28
;    NUMBER OF SEQ ID NOS: 5458
;    SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2758
```

```
;  LENGTH: 384
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-11-293-697-2758

Query Match             3.9%;  Score 90;  DB 7;  Length 384;
Best Local Similarity  24.2%;  Pred. No. 1.2;
Matches   66;  Conservative  24;  Mismatches   95;  Indels  88;  Gaps  15;

Qy      5 LIPPLSLLI----PLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60
            |:|||       |:|  |:|:|  || |||  ||| ||
Db     69 LLPELPLPTQLSSSRPTRPRRCKSSSRPHL-----------ARGRRSPLKAAQ--QDRGLPAWG 117

Qy     61 TWPVKLGWLTPRG--GELIAYLGHY-QRQRLVADGLLAKK---------GCPQSGQVAII 108
            | :||  ||      | ||| || | ||||| ||  ||          |:| |||||
Db    118 VWLLSRAEVPRPSALGSSRRGLGDKGPRSRRAAEFTQARPTRKRRTPCGGRPEAG--CWL 175

Qy    109 ADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILS 168
            :  | :     |||||  | ||         |||  |:|
Db    176 ENRRQRRRKAG-AERAGAGP----TLTPPLPAGSPDPP-SPAATAAPQPEQR----PHLS 225

Qy    169 RAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREKQDE-----SCSLTQALPSELKVS 223
            | |                                       ||||  |||||||||
Db    226 RG------------------------------QWRLLPRRPERAAGSAGSCSSRRALPA----S 256

Qy    224 ADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGW 256
            ||
Db    257 GSGAAMTGSWA---------------------APPRPAW 274
```

```
RESULT 11
US-11-293-697-3284
;  Sequence 3284, Application US/11293697
;  Publication No. US2006010537GA1
;  GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: H1-A0106
;  CURRENT APPLICATION NUMBER: US/11/293,697
;  CURRENT FILING DATE: 2005-12-05
;  PRIOR APPLICATION NUMBER: US/10/108,260
;  PRIOR FILING DATE: 2002-03-28
;  NUMBER OF SEQ ID NOS: 5458
;  SOFTWARE: PatentIn Ver. 2.1
;  SEQ ID NO 3284
;  LENGTH: 952
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-11-293-697-3284

Query Match             3.8%;  Score 87.5;  DB 7;  Length 952;
Best Local Similarity  21.7%;  Pred. No. 7.8;
Matches   84;  Conservative  50;  Mismatches  144;  Indels 109;  Gaps  19;

Qy      8 FLSLL--IPLTPQSAFAQSEPE-LKLESVVIVSRHG-------VRAPTKATQLMQDVTPDA  58
            | ||     |:||| |  ||| ||  ||| ||||         ||  ||
Db    123 FSSLLSDVSILPNRAEKAGGESVQRTSVSAAKEHRRRMGKLLRRASK-QNIRRDCNPEV 181

Qy     59 WPTWPVKLGWLTP------RGGELIA-YLGHYQ-------RQRLVA--------DGLL  94
            ||  |  ||||           ||||  |||||        ||||         ||||
Db    182 SESFKVKDEVFVPFTKPWDCGAELAATSINHKQNHLSLSVEHQSLQACDDLLPYPHGLL 241

Qy     95 AKK-------------GCPQSGQVAIIADVDERTRK----------TGEAFAAGLAPDCAI 132
            |||             |:| |||||||||  |||:          |||| | | |: |
Db    242 SKQWPCSSMPAKQAPPSCVSEG------SVKGRTQKENLFQTNKLKSLSRLAGKGPD--- 292

Qy    133 TVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGHRQTAFRELE-RV 191
            |||||||||||||           || || ||:| |||| |||||||| |||| ||  ||
Db    293 SFEMEEVQSFEEETGNPLDM----TSGTVGARVDRANSCQSDSSGFLEEPLEPLPLQM 346

Qy    192 LNFPQSNLCLK--REKQDESCSLTQALPSELK--------------VSADNVSLTGAVSL 235
            ||||| ||                 ||||| |               |||||||||| |
Db    347 PSLPNSQSPAENGGRKPRDQSHSLVSSQDCQLESDGPDSKSRASMSFSSQEVNALEQRAS 406
```

```
Qy    236 ASMLTEIFLLQQAQGMPE--------------------PGWGRITDSHQWNTLLSLHNAQFYLLQRT 282
            |   | ||    | ||                       :|:  :|:
Db    407 VSVMEBEFLLEAMEGPPELYIPDMACAKTTTRGBCPRKODSHLWQ-LLPMPHAEY------ 459

Qy    283 PEVARSRATPLLDLIMAALTPHPPQKQ 309
            |::| :|| |  :|
Db    460 -EVTRPTATSKYDHPLGFMVTHVTEMQ 485
```

```
RESULT 12
US-10-511-937-2564
;  Sequence 2564, Application US/10511937
;  Publication No. US2006008836A1
;  GENERAL INFORMATION:
;  APPLICANT: EXPRESSION DIAGNOSTICS, INC.
;  APPLICANT: Wohlgemuth, Jay
;  APPLICANT: Fry, Kirk
;  APPLICANT: Woodward, Robert
;  APPLICANT: Ly, Ngoc
;  APPLICANT: Prentice, James
;  APPLICANT: Morris, MacDonald
;  APPLICANT: Rosenberg, Steven
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND MONITORING TRANSPLANT REJECTION
;  FILE REFERENCE: 50661200010A
;  CURRENT APPLICATION NUMBER: US/10/511,937
;  CURRENT FILING DATE: 2004-10-19
;  PRIOR APPLICATION NUMBER: PCT/US2003/012946
;  PRIOR FILING DATE: 2003-04-24
;  PRIOR APPLICATION NUMBER: US 10/131,831
;  PRIOR FILING DATE: 2002-04-24
;  PRIOR APPLICATION NUMBER: US 10/325,899
;  PRIOR FILING DATE: 2002-12-20
;  NUMBER OF SEQ ID NOS: 3117
;  SOFTWARE: PatentIn version 3.2
;  SEQ ID NO 2564
;  LENGTH: 1336
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-10-511-937-2564

Query Match             3.7%;  Score 86;  DB 6;  Length 1336;
Best Local Similarity  21.4%;  Pred. No. 17;
Matches   99;  Conservative  56;  Mismatches 163;  Indels 144;  Gaps  26;

Qy     12 LIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP-----TKATQLMQDVTPDAWPTWPVKL  66
            ||| |             | | ||||  |                :||
Db    773 LASLTPGTGQALRGP-LYTSLAQLYSHHGCHGPAITFMTQAVEASAIAGVRAIVDHLVAL 831

Qy     67 GWLTPRGGELIAYLGHYQ--RQRLVA----DGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
            |  :   ||||||||||    |||||      |:::
Db    832 AWLHVLHGQSPVALDILQSVRDAVVASEDQEGVIA------NMVAVALKRTGRTRQAAE 884

Qy    121 AP------------------AAGLAPDCAITVHTQADTSSPD-----PLFNPLKTYGVCQLD 158
                                 | | || ||
Db    885 SYYRALRVARDLGQORNQAVGLANFGALCLHAGASRLAQHYLLEAVRLFSRLPLGECGRD 944

Qy    159 NANV---TDAILSRAGGS-------IADFTGHRQTAFRELERVLNF-----PQSN 198
            |||     ||||||:||          | | | || ||||
Db    945 FTHVLLQGHLCTRQGPAQQGKGYEWALLVAVEMGHVBSQLRAVQRLCHFYSAVMPSBA 1004

Qy    199 LC-LKREKQ-DBSCSLT-QALPSELKVSADNVSLT------------GAVSLASMLTEIFL-LQ 246
            ||  |:|:|   ||:|   ||  |||||| ||            || | |||||  | ||
Db    1005 QCVIYHELQLSPACKVADKVLEQLLETISQLYLSLGTERAYKSALDYTKRSLGIFIDLQ 1064

Qy    247 QAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAALTPHPP 306
            ||| |   || |||:|||| ||| | |||||
Db    1065 KKE------KEAHAW----------LQAGKIYYILRQSE------------LVDLYI-----  1093

Qy    307 QKQAYGVTLPTSVLPIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFE-RWRRLSD 365
            |:|                           | | |             |||
Db    1094 -QVAQNVALYTG--------DPNL-------GLEL---------------FEAAGDIFFDGAWER--- 1127
```

```
Qy    366  NSQWIQVSLVFQTLQQMRDKT-PLSLNTPPGEVKLTLAGCEE 406
                ||:|| ||::|  |||:| ||| ||
Db   1128  ------------EKAVSFYRDRALPLAVTT--GNRKAELRLCNK 1157
```

```
Db    253  ------IRLLPRAQDGVDDGFWRGEPGPGRVGVFPSLLVEBLLGPPGPPE-------LS 297

Qy    295  DLIMAALTPHPQKQAYGVTLPTSVL 320
                ||:|
Db    298  DPEQMLPSPSPP---SFSSPAPTSVL 320
```

RESULT 13
US-10-196-749-350
; Sequence 350, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Chen,Jian
; APPLICANT: Desnoyers,Luc
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 350
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-350

Query Match   3.6%; Score 84; DB 6; Length 370;
Best Local Similarity   22.9%; Pred. No. 4.1;
Matches   61; Conservative 38; Mismatches 123; Indels 44; Gaps 12;

Qy   68 WLTPRGGELIAVLGHYQRQRLVADGLLAKKKGCPQSGQVAIIADVDERTRKTGEAFAAGLA 127
              ::|| | ||:| |:|||:| :| ||
Db   86 WLKP--AMTQAQDEVEQERRLSEARLSQRDLSPTAEDAELSDFEE-CEETGBLFEEP-A 140

Qy   128 PDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDA----ILSRAGGS--IADFTGHR 181
              |||: :| :| ||: ||::|: : |||||||| ||| |:||
Db   141 -PQALAT-----RALPCPAHVVFRYQAGREDELTITBGEWLEVIBEGDADEWVKARNQHG 194

Qy   182 QTAFRELERVLNFPQSNLCLKREKQDESCS---------LTQALPSELKVSADNVSL-TGAVS 234
              |||||||||| |||| ::||||||| ||||||||||||| ||:||
Db   195 EVGFVP-ERYLNFPDLSLPESSQDSDNPCGABPTAPLAQALYSYTGQSAEElSFPEGAL- 252

Qy   235 LASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLL 294
              ||||| |||||||||||||| ||||:|||||||||||||||||||||||||:
```

RESULT 14
US-11-293-697-3258
; Sequence 3258, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: H1-A0106
;  CURRENT APPLICATION NUMBER: US/11/293,697
;  CURRENT FILING DATE: 2005-12-05
;  PRIOR APPLICATION NUMBER: US/10/108,260
;  PRIOR FILING DATE: 2002-03-28
;  NUMBER OF SEQ ID NOS: 5458
;  SOFTWARE: PatentIn Ver. 2.1
;  SEQ ID NO 3258
;  LENGTH: 725;
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-11-293-697-3258

Query Match         3.6%;  Score 84;  DB 7;  Length 725;
Best Local Similarity   19.9%;  Pred. No. 11;
Matches   68;  Conservative  52;  Mismatches  92;  Indels 130;  Gaps  19;

Qy   119  GEAFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFT 178
              |:  ||| ||: :|||||:|            |  ||:: :| :| ||:
Db    23  GQESTPGIVPD-AVQVHTEEQQKS-----HPSES-----LD----TDSLL-----IAVSS 62

Qy   179  GHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASM 238
              |     :| ::|:|                |:::|          DLKMSSEVKHICGEDQIEDK 101
Db    63  QH--TVNTELEKQIS----------------NEVDSE-----

Qy   239  LTEIFLLQQAQGMPEPGWGRITD-----SHQWNTLLSLHNAQFYLLQRTPEVARSRATPL 293
              :                    ||:           |||::
Db   102  M-----------------EVTENIEVVTHQ---ITVQQEQLQLLEB-PETVVSRE--- 135

Qy   294  LDLIMAALTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGG 353
              ||:|||||||||||||||:||: |||:         |: |: :|:
Db   136  ------------ESRPPKLVMESVTLPLETL-VSPHESI------SLCPEE 168

Qy   354  ELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGC-EERNAQGM 412
              :| |||                                  ||||| :  ::
Db   169  QLVIER-----------------------LQGEKEQKENSELSTGLMDSEMTPTIEGCVKDVSYQGG 212

Qy   413  CSL-----------AGFTQI--------VNEARIPACSLRSHHHHHH 440
              : |                |||| |||
Db   213  KSIKLSSETESSFSSSADISKADVSSSPTPSSDLPSHDMLHN 254
```

RESULT 15
US-10-953-349-22123
; Sequence 22123, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCONDED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22123
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Glycine max

Thu Jun 15 11:38:59 2006

us-10-601-319-2.rapbn

US-10-953-349-22123

```
Query Match          3.6%;     Score 82.5;   DB 6;   Length 360;
Best Local Similarity  21.1%;    Pred. No. 5.4;
Matches   82;  Conservative  53;  Mismatches  155;  Indels  99;  Gaps  18;
```

```
Qy   34 VIVSRHGVRAPTKATQLMQDVTPDAWPTWPVKLGWLTPRGGE-----LIAYLGHY----Q  84
           :; ::||  |:  |                      :  : :  |:::  |
Db    6 IAIKRHASDLTTKSMALVA-----------LLKLSSRFPSCSERIKEIIVQFKGSFVLELQ  55

Qy   85 RQRIVADGLLAKKGCPQSGQVAIIADVDERT---RKTGEAFAAGLAPDCAITVHTQADTS 141
           :: :  ::||   :: || ||  |  |        :   |:
Db   56 QRAIEFNSIIAKHQNIRSTLVERMPVLDEATSIGRRAGSLPGAASTP-------TAPSFN 108

Qy  142 SPDPLFNPLKTGVCQLD-NANVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLC 200
           |:                  ||  :::                  |  :|
Db  109 LPNGTAKPVAPLVDLLDLSSDDAPAPSSSSGGDI-----------LQDLLGV------- 149

Qy  201 LKREKQDESCSLTQALPSELKVSADNVSL----TGAVSLASMLTEIFLLQQAQGMPEPGW 256
           |              ::: :  :: |  |    | :   |: ::  :
Db  150 ------DLSPASQQSVAGQASKSGNDVLLDLLSIGSPSVESSSSTVDILSSNSSNKAP-- 201

Qy  257 GRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAALTPHPPQKQAYGVTLP 316
           |               |: : :              | ::::  |  |  |  :  |
Db  202 ---VSSLDGLSSLSLSTK----------TTSNAAPMMDL-LDGFAPIPPTEN-NGPVYP 245

Qy  317 TSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVF 376
           :       :  ::||| |                  : :|  |            ||
Db  246 SVTAFES---------SSLRLTFNFSKQPGN--PQTTVIQATFMNLSSNT---YTDFVF 290

Qy  377 QT-----LQQMRDKTPLSLNTPPGEVKLT 400
           |      :   |:::     |: :|
Db  291 QAAVPKFQLHLD--PASSNTLPANGSIT 317
```

Search completed: June 13, 2006, 10:51:02
Job time : 12.0917 secs

THIS PAGE LEFT BLANK

```
                 GenCore version 5.1.9
          Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on:         June 13, 2006, 10:17:14 ; Search time 139.266 Seconds
                                           (without alignments)
                                  1444.539 Million cell updates/sec

Title:          US-10-601-319-2
Perfect score:  2302
Sequence:       1 MKAILIPFLSLLIPLTPQSA.........IVNEARIPACSLRSHHHHHH 440

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters:     2589679

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      A_Geneseq_8:*
                1: geneseqp1980s:*
                2: geneseqp1990s:*
                3: geneseqp2000s:*
                4: geneseqp2001s:*
                5: geneseqp2002s:*
                6: geneseqp2003as:*
                7: geneseqp2003bs:*
                8: geneseqp2004s:*
                9: geneseqp2005s:*
               10: geneseqp2006s:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2302	100.0	440	4	AAB37892	Aab37892 Escherich
2	2302	100.0	440	5	AAE22836	Aae22836 Escherich
3	2302	100.0	440	5	AAE15806	Aae15806 Escherich
4	2302	100.0	440	6	ADA19446	Ada19446 E. coli B
5	2302	100.0	440	8	ADO50292	Ado50292 Escherich
6	2302	100.0	440	10	AEE75413	Aee75413 B. coli B
7	2282	99.1	440	8	ADO50304	Ado50304 Escherich
8	2279	99.0	440	8	ADO50302	Ado50302 Kangaroo
9	2235	97.1	432	3	AAB36257	Aab36257 Lama2/APP
10	2235	97.1	432	3	AAB36261	Aab36261 R15/APPA
11	2235	97.1	432	3	AAB36262	Aab36262 SV40/APPA
12	2235	97.1	432	3	AAB36259	Aab36259 R15/APPA
13	2235	97.1	432	3	AAB36258	Aab36258 R15/APPA
14	2235	97.1	432	3	AAB36263	Aab36263 Lama2/APP
15	2235	97.1	432	3	AAB36260	Aab36260 R15/APPA
16	2235	97.1	432	3	AAU77775	Aau77775 Phytase a
17	2235	97.1	432	5	AAE15807	Aae15807 Escherich
18	2235	97.1	432	6	ABP53625	Abp53625 Acid phos
19	2235	97.1	432	6	ADA19450	Ada19450 E. coli K
20	2235	97.1	432	6	ABU15427	Abu15427 Protein e
21	2235	97.1	432	7	ADC87745	Adc87745 Escherich
22	2235	97.1	432	8	ADL16129	Adl16129 Escherich
23	2235	97.1	432	8	ADO50298	Ado50298 Escherich
24	2235	97.1	432	9	AED50820	Aed50820 Escherich
25	2235	97.1	432	10	AEE75419	Aee75419 E. coli B
26	2232	97.0	432	10	AEF15220	Aef15220 Escherich
27	2230	96.9	432	4	AAE02631	Aae02631 E. coli a
28	2230	96.6	432	8	ADL16131	Adl16131 Shigella
29	2224	96.6	432	8	AAE02634	Aae02634 E. coli a
30	2224	96.6	432	8	ADL16130	Adl16130 Shigella
31	2223	96.6	432	3	AAE02635	Aae02635 E. coli a
32	2221	96.5	432	3	AAY94753	Aay94753 E. coli a
33	2221	96.5	432	6	ABR42161	Abr42161 Escherich
34	2210	96.0	432	7	AAE02632	Aae02632 E. coli a
35	2198	95.5	432	7	ADL02198	Adl02198 AppA. 5/2
36	2197	95.4	432	6	ABR42162	Abr42162 Escherich
37	2194.5	95.3	423	7	AAY01513	Aay01513 An Escher
38	2188	95.0	436	7	ADC87743	Adc87743 Modified
39	2182	94.8	432	10	AEF15219	Aef15219 Escherich
40	2171.5	94.3	431	8	ADW76361	Adw76361 Phytase A
41	2171	94.3	432	8	ADO50300	Ado50300 Escherich
42	2171	94.3	432	9	AED50818	Aed50818 Escherich
43	2171	94.3	432	10	AEE75421	Aee75421 E. coli B
44	2141	93.0	430	5	AAE15808	Aae15808 Escherich
45	2141	93.0	430	6	ADA19451	Ada19451 E. coli K

ALIGNMENTS

```
RESULT 1
AAB37892
ID   AAB37892 standard; protein; 440 AA.
XX
AC   AAB37892;
XX
DT   07-MAR-2001  (first entry)
XX
DE   Escherichia coli B phytase enzyme.
XX
KW   Escherichia coli B; phytase enzyme; anabolic; phytate digestion;
KW   nutrition.
XX
OS   Escherichia coli.
XX
PN   WO20071728-A1.
XX
PD   30-NOV-2000.
XX
PF   25-MAY-2000; 2000WO-US014846.
XX
PR   25-MAY-1999;  99US-00318528.
XX
PA   (DIVE-) DIVERSA CORP.
XX
PI   Short JM, Kretz KA;
XX
DR   WPI; 2001-112081/12.
DR   N-PSDB; AAC88885.
XX
PT   Improving the nutritional value of phytate-containing foodstuffs, using
PT   phytase enzymes which catalyze the liberation of inorganic phosphate from
PT   the phytates.
XX
PS   Claim 1; Fig 1; 147pp; English.
XX
CC   The present sequence is a phytase enzyme from Escherichia coli B. The
CC   enzyme catalyses the liberation of inorganic phosphate from the phytate
CC   in phytate-containing foodstuffs and can thus be used to improve the
CC   nutritional value of phytate rich ingredients
XX
SQ   Sequence 440 AA;

Query Match             100.0%;  Score 2302;  DB 4;  Length 440;
Best Local Similarity   100.0%;  Pred. No. 2.7e-223;
Matches  440;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy  181 RQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy  421 IVNEARIPACSLRSHHHHHH 440
        ||||||||||||||||||||
Db  421 IVNEARIPACSLRSHHHHHH 440

RESULT 2
AAE22836
ID   AAE22836 standard; protein; 440 AA.
XX
AC   AAE22836;
XX
DT   21-AUG-2002  (first entry)
XX
DE   Escherichia coli phytase protein.
XX
KW   Dietary aid; biocompatible composition; therapeutic; digestive tract;
KW   foodstuff; digestion; phytase; enzyme.
XX
OS   Escherichia coli.
XX
PN   WO200189317-A2.
XX
PD   29-NOV-2001.
XX
PF   15-MAY-2001; 2001WO-US015764.
XX
PR   25-MAY-2000; 2000US-00580937.
XX
PA   (DIVE-) DIVERSA CORP.
XX
PI   Short JM, Kretz KA, O'donoghue E;
XX
DR   WPI; 2002-164119/21.
DR   N-PSDB; AAD36473.
XX
PT   New dietary aids comprising sustained release biocompatible compositions,
PT   comprise agent that assists in digestion, useful for delivering enzymes,
PT   therapeutics, medicine or agents to an organism.
XX
PS   Claim 5; Fig 1; 89pp; English.
XX
CC   The present invention relates to novel dietary aids comprising sustained
CC   release biocompatible composition which comprises an agent (enzymes such
CC   as phytase, amylase, esterase, protease) that assists in digestion. The
CC   biocompatible composition is effective upon oral consumption and release
CC   in the digestive tract of a subject. The dietary aids are useful for
CC   delivering enzymes, therapeutics, medicine and agents to an organism. The
CC   use of enzymes and other agents in digestive aids of livestock or
CC   domesticated animals not only improves the animal's health and life
CC   expectancy but also assists in increasing the health of livestock or in
CC   the production of foodstuffs from livestock. The present sequence is
CC   Escherichia coli phytase protein
XX
SQ   Sequence 440 AA;

Query Match             100.0%;  Score 2302;  DB 5;  Length 440;
Best Local Similarity   100.0%;  Pred. No. 2.7e-223;
Matches  440;  Conservative   0;  Mismatches   0;  Indels    0;  Gaps    0;

Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy  181 RQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy  421 IVNEARIPACSLRSHHHHHH 440
        ||||||||||||||||||||
Db  421 IVNEARIPACSLRSHHHHHH 440

RESULT 3
AAE15806
ID   AAE15806 standard; protein; 440 AA.
XX
AC   AAE15806;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Escherichia coli B phytase protein.
XX
KW   Bacterial phytase; K12 appA phytase; protease stability; anabolic;
KW   gastrointestinal; nutritional value; feed treatment process; therapy;
KW   thermal tolerance; growth performance; alcoholic drink; biopulping;
KW   non-alcoholic drink; biobleaching; B phytase.
XX
OS   Escherichia coli.
XX
PN   WO200190333-A2.
XX
PD   29-NOV-2001.
XX
PF   24-MAY-2001; 2001WO-US017118.
XX
PR   25-MAY-2000; 2000US-00580515.
XX
```

```
PA   (DIVE-) DIVERSA CORP.
XX
PI   Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, O'donoghue E;
XX
DR   WPI; 2002-083108/11.
DR   N-PSDB; AAD25460.
XX
PT   New bacterial phytase for e.g. improving the nutritional value of phytate
PT   -containing foodstuffs and subsequently improving the growth performance
PT   of an organism that consumes it, or in treating animal digestive systems.
XX
PS   Claim 1; Fig 1; 170pp; English.
XX
CC   The patent discloses recombinant bacterial phytase from Escherichia coli
CC   K12 appA phytase. The enzyme has phytase activity and improved thermal
CC   tolerance when compared with wild-type phytase. It has improved protease
CC   stability at low pH. The recombinant phytase is useful for improving the
CC   nutritional value of phytate-containing foodstuffs and subsequently
CC   improving the growth performance of an organism that consumes it, in
CC   treating animal digestive systems, in feed treatment processes and for in
CC   vitro purposes related to research, discovery and development. They are
CC   also used for generating recombinant digestive system life forms, for
CC   producing or manufacturing alcoholic and non-alcoholic drinks based on
CC   the use of moulds, grains and/or plants, in biopulping and bio-bleaching
CC   where a reduction in the use of environmentally harmful chemicals that
CC   are traditionally used in the pulp and paper industry is desired and in
CC   the reduction or possible elimination of the need for mineral
CC   supplements, enzymes or therapeutic drugs for animals from the daily feed
CC   thus increasing the amount calories and nutrients present in the feed.
CC   The present sequence is E. coli B phytase protein
XX
SQ   Sequence 440 AA;

Query Match             100.0%;  Score 2302; DB 5; Length 440;
Best Local Similarity   100.0%;  Pred. No. 2.7e-223;
Matches 440;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60

Qy     61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120

Qy    121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADPTGH  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADPTGH  180

Qy    181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240

Qy    241 BIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 BIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300

Qy    301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360

Qy    361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKJTLAGCEERNAQGMCSLAGPTQ  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKJTLAGCEERNAQGMCSLAGPTQ  420

Qy    421 IVNEARIPACSLRSHHHHHH  440
          ||||||||||||||||||||
Db    421 IVNEARIPACSLRSHHHHHH  440


RESULT 4
ADA19446
ID   ADA19446 standard; protein; 440 AA.
XX
AC   ADA19446;
XX
DT   20-NOV-2003  (first entry)
XX
DE   E. coli B phytase.
XX
KW   Phytase; enzyme; phytate; appA gene; animal feed; inorganic phosphate;
KW   digestion enhancement; transgenic; thermal tolerance; protease stability.
XX
OS   Escherichia coli B.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 72
FT                   /note= "Encoded by CGN"
XX
PN   US2002136754-A1.
XX
PD   26-SEP-2002.
XX
PF   24-MAY-2001; 2001US-0086379.
XX
PR   13-AUG-1997;  97US-00910798.
PR   01-MAR-1999;  99US-00259214.
PR   13-APR-1999;  99US-00291931.
PR   25-MAY-1999;  99US-0031528.
PR   25-MAY-2000;  2000US-0580515.
XX
PA   (SHOR/)  SHORT J M.
PA   (KRET/)  KRETZ K A.
PA   (GRAY/)  GRAY K A.
PA   (BART/)  BARTON N R.
PA   (GARR/)  GARRETT J B.
PA   (DONO/)  O' DONOGHUE E.
PA   (MATH/)  MATHUR E J.
XX
PI   Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB;
PI   O'Donoghue E, Mathur EJ;
XX
DR   WPI; 2003-040002/03.
XX
PT   Isolated Escherichia coli polynucleotide encoding a modified phytase
PT   enzyme, useful in the production of animal feed, for improving nutritional
PT   value of phytate-containing foodstuff and subsequently
PT   digestion in humans and animals.
XX
PS   Claim 1; Fig 1; 62pp; English.
XX
CC   The invention relates to an isolated Escherichia coli polynucleotide
CC   encoding a phytase enzyme appearing as ADA19450 and having amino acids
CC   modification from W68B, Q84W, A95P, K97C, S168E, R180Y, N226C or Y277D.
CC   Also included the E. coli appA gene ADA19449 (or an oligonucleotide
CC   derived from it) or its mutant sequence ADA19452, expression vectors,
CC   host cells, a method of improving nutritional value of a phytate-
CC   containing foodstuff by contacting the phytate-containing foodstuff with
CC   a pure phytase mutated phytase (where the phytase enzyme catalyzes
CC   the liberation of inorganic phosphate from the phytate in the phytate-
CC   containing foodstuff), a method to produce an animal feed containing a
CC   microbial phytase (comprising culturing the plant cell, plant part or
CC   plant under conditions where the nucleotide sequence is expressed and
CC   converting the plant cells, plants or plant into a composition for animal
CC   feed), a feed composition for animals (comprising the plant seeds, plant
CC   cells, plant parts or plants in admixture with a phytate-containing
CC   foodstuff), a method to treat a human or an animal able to benefit from
CC   digestive enhancement by the activity of an exogenous phytase enzymes
CC   comprising administering to the human or animal the plant seed, plant
CC   cells, plant parts or plants of a transgenic plant which is modified to
CC   contain an expression system which expresses a nucleotide sequence
CC   encoding a phytase enzyme, a transgenic non-human organism whose genome
CC   comprising a heterologous nucleic acid sequence encoding a polypeptide
CC   having phytase activity. The phytate enzyme is useful for improving the
CC   nutritional value of phytate-containing foodstuff, in the production of
CC   animal feed and for enhancing digestion in humans and animals. The
```

```
CC   invented method improves thermal tolerance and protease stability. It
CC   also improves the feeding value of phytate rich ingredients. The present
CC   sequence represents E. coli B wild-type phytase.
XX
SQ   Sequence 440 AA;

Query Match             100.0%;  Score 2302;  DB 6;  Length 440;
Best Local Similarity   100.0%;  Pred. No. 2.7e-223;
Matches  440;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP  60

Qy     61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy    121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy    181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy    241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy    301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy    361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy    421  IVNEARIPACSLRSHHHHHH 440
           ||||||||||||||||||||
Db    421  IVNEARIPACSLRSHHHHHH 440

RESULT 5
ADO50292
ID   ADO50292 standard; protein; 440 AA.
XX
AC   ADO50292;
XX
DT   29-JUL-2004  (first entry)
DE   Escherichia coli B phytase enzyme.
XX
KW   Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW   phytate; animal feed; fish feed; dough; baking; enzyme.
XX
OS   Escherichia coli; B.
XX
FH   Key            Location/Qualifiers
FT   Misc-difference 72
FT                  /note= "Encoded by CGN"
XX
PN   US2004091968-A1.
XX
PD   13-MAY-2004.
XX
PR   13-AUG-1997;  97US-00910798.
PR   01-MAR-1999;  99US-00259214.
PR   13-APR-1999;  99US-00291931.
PR   25-MAY-1999;  99US-00318528.
PR   25-MAY-2000;  2000US-00580515.


PR   24-MAY-2001;  2001US-00866379.
XX
PA   (SHOR/)  SHORT J M.
PA   (KRET/)  KRETZ K.
PA   (GRAY/)  GRAY K A.
PA   (BART/)  BARTON N R.
PA   (GARR/)  GARRETT J B.
PA   (ODON//) O'DONOGHUE E.
PA   (MATH//) MATHER E J.
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI   Mather EJ;
XX
DR   WPI; 2004-374952/35.
DR   N-PSDB; ADO50291.
XX
PT   Producing phytase, involves providing nucleic acid derived from bacteria
PT   encoding polypeptide having phytase activity, and expressing nucleic acid
PT   in yeast.
XX
PS   Claim 3; SEQ ID NO 2; 74pp; English.
XX
CC   The invention relates to a method for producing phytase that involves
CC   providing a nucleic acid encoding phytase derived from a bacteria, and
CC   expressing the nucleic acid in a yeast under conditions that allow
CC   expression of the enzyme in the yeast. The invention also relates to
CC   modified phytase enzyme which has improved thermal tolerance and protease
CC   stability at low pH. The phytase enzyme can be used in foodstuffs to
CC   improve the feeding value of phytate rich ingredients, and in diet of
CC   numerous animals including mammals, fowls and fishes, commercially
CC   significant mammals such as pigs, goats, laboratory rodents, commercially
CC   significant avian species such as chicken, ducks, doves, parrot, etc.,
CC   commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC   etc., in dough making and baking, in dietary aids for animals. The method
CC   provides easy manufacture of the active ingredient loaded biocompatible
CC   composition, higher yields and loading efficiency. The phytase
CC   incorporated in to the dietary aid is safe for animals. The present
CC   sequence is Escherichia coli B phytase enzyme.
XX
SQ   Sequence 440 AA;

Query Match             100.0%;  Score 2302;  DB 8;  Length 440;
Best Local Similarity   100.0%;  Pred. No. 2.7e-223;
Matches  440;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Sequence 440 AA;

Qy      1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP  60

Qy     61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy    121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy    181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy    241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy    301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy    361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
```

```
Qy    421 IVNEARIPACSLRSHHHHHH 440
          ||||||||||||||||||||
Db    421 IVNEARIPACSLRSHHHHHH 440

RESULT 6
AEE75413
ID   AEE75413 standard; protein; 440 AA.
XX
AC   AEE75413;
XX
DT   23-FEB-2006  (first entry)
XX
DE   E. coli B phytase SEQ ID NO: 2.
XX
KW   enzyme; thermostable; phytase; protein stabilization; pharmaceutical;
KW   fertilizer; cramp; cell culture; osteoporosis; osteopathic; food.
XX
OS   Escherichia coli B.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 72
FT                   /note= "Encoded by CGN"
XX
PN   US2005281792-A1.
XX
PD   22-DEC-2005.
XX
PF   01-SEP-2004; 2004US-00933115.
XX
PR   13-AUG-1997;   97US-00910798.
PR   01-MAR-1999;   99US-00259214.
PR   13-APR-1999;   99US-00291931.
PR   25-MAY-1999;   99US-00318528.
PR   25-MAY-2000; 2000US-00580515.
PR   24-MAY-2001; 2001US-00866379.
XX
PA   (SHOR/)  SHORT J M.
PA   (KRET/)  KRETZ K A.
PA   (GRAY/)  GRAY K A.
PA   (BART/)  BARTON N R.
PA   (GARR/)  GARRETT J B.
PA   (ODON/)  ODONOGHUE B.
PA   (BAUM/)  BAUM W.
PA   (ROBE/)  ROBERTSON D B.
PA   (ZORN/)  ZORNER P.
XX
PI   Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, Odonoghue E;
PI   Baum W, Robertson DE, Zorner P;
XX
DR   WPI; 2006-055022/06.
DR   N-PSDB; AEE75412.
XX
PT   Formulation useful as dietary supplement for treating, preventing or
PT   reversing osteoporosis and bone loss, and preventing muscle cramps,
PT   comprises one or more polypeptides having phytase activity.
XX
PS   Claim 1; SEQ ID NO 2; 82pp; English.
XX
CC   The present sequence is that of the Escherichia coli B phytase. The
CC   present invention relates to a novel formulation, comprising at least one
CC   polypeptide having phytase activity, where the polypeptide is a modified
CC   sequence (AEE75421) derived from the Escherichia coli K-12 appA phytase
CC   (AEE75419). The modification of the enzyme, comprising a series of
CC   mutations, improves the thermal tolerance and protease stability of the
CC   protein. The specification also claims a pharmaceutical composition; a
CC   kit; an immobilized phytase; a fertilizer or soil additive; a liquid
CC   supplement for preventing muscle cramps; a hydrating agent; a tissue
CC   culture or cell culture media; and a plant food additive, all comprising
CC   the novel phytase, and a method of reducing pollution and increasing
CC   nutrient availability in an environment or environmental sample by
CC   degrading environmental phytic acid. The novel enzyme is a dietary
CC   supplement useful for treating, preventing or reversing osteoporosis or
CC   bone loss, and preventing muscle cramps. The liquid supplement is useful
CC   for preventing muscle cramps. The method is useful for reducing pollution
CC   and increasing nutrient availability in an environment or environmental
CC   sample by degrading environmental phytic acid, where the environment or
CC   environmental sample comprises a soil or a body of water. The immobilized
CC   phytase is useful in foodstuffs for improving the feeding value of
CC   phytate rich ingredients.
XX
SQ   Sequence 440 AA;

Query Match             100.0%;  Score 2302;  DB 10;  Length 440;
Best Local Similarity   100.0%;  Pred. No. 2.7e-223;
Matches  440;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy     61 TWPVKLGWLTPRGGELIAVLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TWPVKLGWLTPRGGELIAVLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy    121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy    181 RQTAPRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 RQTAPRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy    241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy    301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360

Qy    361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy    421 IVNEARIPACSLRSHHHHHH 440
          ||||||||||||||||||||
Db    421 IVNEARIPACSLRSHHHHHH 440

RESULT 7
ADO50304
ID   ADO50304 standard; protein; 440 AA.
XX
AC   ADO50304;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Escherichia coli W phytase 875PH2 mutant enzyme.
XX
KW   Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW   phytate; animal feed; fish feed; dough; baking; mutant; mutein.
XX
OS   Escherichia coli.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 72
FT                   /note= "Encoded by CGN"
FT   Misc-difference 160
FT                   /note= "E. coli B phytase Ala replaced with Ser"
FT   Misc-difference 176
FT                   /note= "E. coli B phytase Asp replaced with Gly"
FT   Misc-difference 298
FT                   /note= "E. coli B phytase Met replaced with Lys"
FT   Misc-difference 299
FT                   /note= "E. coli B phytase Ala replaced with Thr"
```

```
Db   181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300
              |||||||||||||||||||||||||||||||||||||||||||||| ||
Db   241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA  300
Qy   301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360
Qy   361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
Qy   421  IVNEARIPACSLRSHHHHHH  440
              ||||||||||||||||||
Db   421  IVNEARIPACSLRSHHHHHH  440
```

RESULT 8
AD050302
ID   AD050302 standard; protein; 440 AA.
XX
AC   AD050302;
XX
DT   29-JUL-2004 (first entry)
XX
DE   Kangaroo rat Escherichia coli phytase 872PH1 mutant enzyme.
XX
KW   Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW   phytate; animal feed; fish feed; dough; baking; mutant; mutein.
XX
OS   Escherichia coli.
XX
FH   Key   Location/Qualifiers
FT   Misc-difference 72
FT   /note= "Encoded by CGN"
FT   Misc-difference 176
FT   /note= "E. coli B phytase Asp replaced with Gly"
FT   Misc-difference 298
FT   /note= "E. coli B phytase Met replaced with Lys"
FT   Misc-difference 299
FT   /note= "E. coli B phytase Ala replaced with Thr"
FT   Misc-difference 312
FT   /note= "E. coli B phytase Gly replaced with Ser"
XX
PN   US2004091968-A1.
XX
PD   13-MAY-2004.
XX
PF   20-JUN-2003; 2003US-00601319.
XX
PR   13-AUG-1997; 97US-00910798.
PR   01-MAR-1999; 99US-00259214.
PR   13-APR-1999; 99US-00291931.
PR   25-MAY-1999; 99US-00318528.
PR   25-MAY-2000; 2000US-00580515.
PR   24-MAY-2001; 2001US-00866379.
XX
PA   (SHOR/) SHORT J M.
PA   (KRET/) KRETZ K.
PA   (GRAY/) GRAY K A.
PA   (BART/) BARTON N R.
PA   (GARR/) GARRETT J B.
PA   (ODON/) O'DONOGHUE E.
PA   (MATH/) MATHER E J.
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI   Mather EJ;
XX
DR   WPI; 2004-374952/35.
DR   N-PSDB; AD050301.
XX

XX
PN   US2004091968-A1.
XX
PD   13-MAY-2004.
XX
PF   20-JUN-2003; 2003US-00601319.
XX
PR   13-AUG-1997; 97US-00910798.
PR   01-MAR-1999; 99US-00259214.
PR   13-APR-1999; 99US-00291931.
PR   25-MAY-1999; 99US-00318528.
PR   25-MAY-2000; 2000US-00580515.
PR   24-MAY-2001; 2001US-00866379.
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI   Mather EJ;
XX
DR   WPI; 2004-374952/35.
DR   N-PSDB; AD050291, AD050303.
XX
PT   Producing phytase, involves providing nucleic acid derived from bacteria
PT   encoding polypeptide having phytase activity, and expressing nucleic acid
PT   in yeast.
XX
PS   Disclosure; Page; 74pp; English.
XX
CC   The invention relates to a method for producing phytase that involves
CC   providing a nucleic acid encoding phytase derived from a bacteria, and
CC   expressing the nucleic acid in a yeast under conditions that allow
CC   expression of the enzyme in the yeast. The invention also relates to
CC   modified phytase enzyme which has improved thermal tolerance and protease
CC   stability at low pH. The phytase enzyme can be used in foodstuffs to
CC   improve the feeding value of phytate rich ingredients, and in diet of
CC   numerous animals including mammals, fowls and fishes, commercially
CC   significant mammals such as pigs, goats, laboratory rodents, commercially
CC   significant avian species such as chicken, ducks, doves, parrot, etc.,
CC   commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC   etc., in dough making and baking, in dietary aids for animals. The method
CC   provides easy manufacture of the active ingredient loaded biocompatible
CC   composition, higher yields and loading efficiency. The phytase
CC   incorporated in to the dietary aid is safe for animals. The present
CC   sequence is Escherichia coli W phytase mutant enzyme. Note: The present
CC   sequence is not shown in the specification but has been derived from
CC   Escherichia coli B phytase AD050292.
XX
SQ   Sequence 440 AA;

Query Match   99.1%; Score 2282; DB 8; Length 440;
Best Local Similarity   99.1%; Pred. No. 2.9e-221;
Matches 436; Conservative   1; Mismatches   3; Indels   0; Gaps   0;

```
Qy     1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60
Qy    61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120
Qy   121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180
             |||||||||||||||||||||||||||||||  |||||| ||||||||||||||:|||||
Db   121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNSNVTDAILSRAGGSIAGFTGH  180
Qy   181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

PT Producing phytase, involves providing nucleic acid derived from bacteria
PT encoding polypeptide having phytase activity, and expressing nucleic acid
PT in yeast.
XX
PS Disclosure; Page; 74pp; English.
XX
CC The invention relates to a method for producing phytase that involves
CC providing a nucleic acid encoding phytase derived from a bacteria, and
CC expressing the nucleic acid in a yeast under conditions that allow
CC expression of the enzyme in the yeast. The invention also relates to
CC modified phytase enzyme which has improved thermal tolerance and protease
CC stability at low pH. The phytase enzyme can be used in foodstuffs to
CC improve the feeding value of phytate rich ingredients, and in diet of
CC numerous animals including mammals, fowls and fishes, commercially
CC significant mammals such as pigs, goats, laboratory rodents, commercially
CC significant avian species such as chicken, ducks, doves, parrot, etc.,
CC commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC etc., in dough making and baking, in dietary aids for animals. The method
CC provides easy manufacture of the active ingredient loaded biocompatible
CC composition, higher yields and loading efficiency. The phytase
CC incorporated in to the dietary aid is safe for animals. The present
CC sequence is Kangaroo rat Escherichia coli phytase mutant enzyme. Note:
CC The present sequence is not shown in the specification but has been
CC derived from Escherichia coli B phytase sequence AD050292.
XX
SQ Sequence 440 AA;

Query Match　　　 99.0%; 　Score 2279; 　DB 8; 　Length 440;
Best Local Similarity 99.1%; 　Pred. No. 5.7e-221;
Matches 436; 　Conservative 0; 　Mismatches 0; 　Indels 4; 　Gaps 0;

```
Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIAGFTGH 180

Qy  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy  241 EIFLLQQAGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy  421 IVNEARIPACSLRSHHHHHH 440
         ||||||||||||||||||||
Db  421 IVNEARIPACSLRSHHHHHH 440
```

RESULT 9
AAB36257
ID   AAB36257 standard; protein; 432 AA.
XX
AC   AAB36257;
XX
DT   12-SEP-2003 (revised)
DT   20-FEB-2001 (first entry)
XX
DE   Lama2/APPA plasmid translated sequence.
XX
KW   Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW   environmental pollution; pig.
XX
OS   Mus musculus.
OS   Escherichia coli.
OS   Chimeric.
XX
PN   WO20006424 7-A1.
XX
PD   02-NOV-2000.
XX
PF   20-APR-2000; 2000WO-CA000430.
XX
PR   23-APR-1999; 99US-0130508P.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   Forsberg CW, Golovan S, Phillips JP;
XX
DR   WPI; 2000-687245/67.
DR   N-PSDB; AAC68294.
XX
PT   Transgenic non-human animal for gastrointestinal tract specific
PT   expression of a protein, preferably phytase, comprises a nucleic acid
PT   sequence including a heterologous transgene construct encoding the
PT   protein.
XX
PS   Disclosure; Fig 5; 152pp; English.
XX
CC   The present invention provides transgenic animals which produce desired
CC   proteins, in this case pigs which expresses phytase in the salivary
CC   gland. Low phytase production levels result in phytate in the diet being
CC   excreted and causing phosphorus contamination in water, as well as
CC   reducing the growth of animals. The invention provides a number of
CC   transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC   on 12-SEP-2003 to standardise OS field)
XX
SQ   Sequence 432 AA;

Query Match　　　 97.1%; 　Score 2235; 　DB 3; 　Length 432;
Best Local Similarity 99.5%; 　Pred. No. 1.6e-216;
Matches 430; 　Conservative 0; 　Mismatches 2; 　Indels 0; 　Gaps 0;

```
Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy  241 EIFLLQQAGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
```

```
Qy   421 IVNEARIPACSL 432
         ||||||||||||
Db   421 IVNEARIPACSL 432

RESULT 10
AAB36261
ID   AAB36261 standard; protein; 432 AA.
XX
AC   AAB36261;
XX
DT   12-SEP-2003 (revised)
DT   20-FEB-2001 (first entry)
XX
DE   R15/APPA plasmid translated sequence.
XX
KW   Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW   environmental pollution; pig.
XX
OS   Rattus sp.
OS   Escherichia coli.
OS   Chimeric.
XX
PN   WO20006424247-A1.
XX
PD   02-NOV-2000.
XX
PF   20-APR-2000; 2000WO-CA000430.
XX
PR   23-APR-1999; 99US-0130508P.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   Forsberg CW, Golovan S, Phillips JP;
XX
DR   WPI; 2000-687245/67.
DR   N-PSDB; AAC68299.
XX
PT   Transgenic non-human animal for gastrointestinal tract specific
PT   expression of a protein, preferably phytase, comprises a nucleic acid
PT   sequence including a heterologous transgene construct encoding the
PT   protein.
XX
PS   Disclosure; Fig 21; 152pp; English.
XX
CC   The present invention provides transgenic animals which produce desired
CC   proteins, in this case pigs which expresses phytase in the salivary
CC   gland. Low phytase production levels result in phytate in the diet being
CC   excreted and causing phosphorus contamination in water, as well as
CC   reducing the growth of animals. The invention provides a number of
CC   transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC   on 12-SEP-2003 to standardise OS field)
XX
SQ   Sequence 432 AA;

Query Match          97.1%;  Score 2235;  DB 3;  Length 432;
Best Local Similarity 99.5%;  Pred. No. 1.6e-216;
Matches 430;  Conservative 0;  Mismatches 2;  Indels 0;  Gaps 0;

Qy     1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy    61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy   121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy   241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGELVFERW 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGELVFERW 360

Qy   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy   421 IVNEARIPACSL 432
         ||||||||||||
Db   421 IVNEARIPACSL 432

RESULT 11
AAB36262
ID   AAB36262 standard; protein; 432 AA.
XX
AC   AAB36262;
XX
DT   12-SEP-2003 (revised)
DT   20-FEB-2001 (first entry)
XX
DE   SV40/APPA plasmid translated sequence.
XX
KW   Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW   environmental pollution; pig.
XX
OS   Simian virus 40.
OS   Escherichia coli.
OS   Chimeric.
XX
PN   WO20006424247-A1.
XX
PD   02-NOV-2000.
XX
PF   20-APR-2000; 2000WO-CA000430.
XX
PR   23-APR-1999; 99US-0130508P.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   Forsberg CW, Golovan S, Phillips JP;
XX
DR   WPI; 2000-687245/67.
DR   N-PSDB; AAC68299.
XX
PT   Transgenic non-human animal for gastrointestinal tract specific
PT   expression of a protein, preferably phytase, comprises a nucleic acid
PT   sequence including a heterologous transgene construct encoding the
PT   protein.
XX
PS   Disclosure; Fig 22; 152pp; English.
XX
CC   The present invention provides transgenic animals which produce desired
CC   proteins, in this case pigs which expresses phytase in the salivary
CC   gland. Low phytase production levels result in phytate in the diet being
CC   excreted and causing phosphorus contamination in water, as well as
CC   reducing the growth of animals. The invention provides a number of
CC   transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC   on 12-SEP-2003 to standardise OS field)
XX
SQ   Sequence 432 AA;

Query Match          97.1%;  Score 2235;  DB 3;  Length 432;
Best Local Similarity 99.5%;  Pred. No. 1.6e-216;
Matches 430;  Conservative 0;  Mismatches 2;  Indels 0;  Gaps 0;
```

CC   The present invention provides transgenic animals which produce desired
CC   proteins, in this case pigs which expresses phytase in the salivary
CC   gland. Low phytase production levels result in phytate in the diet being
CC   excreted and causing phosphorus contamination in water, as well as
CC   reducing the growth of animals. The invention provides a number of
CC   transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC   on 12-SEP-2003 to standardise OS field)
XX
SQ   Sequence 432 AA;

Query Match   97.1%; Score 2235; DB 3; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.6e-216;
Matches 430;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

```
Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy  421 IVNEARIPACSL 432
          ||||||||||||
Db  421 IVNEARIPACSL 432
```

RESULT 13
AAB36258
ID   AAB36258 standard; protein; 432 AA.
XX
AC   AAB36258;
XX
DT   12-SEP-2003   (revised)
DT   20-FEB-2001   (first entry)
XX
DE   R15/APPA plasmid translated sequence.
XX
KW   Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW   environmental polution; pig.
XX
OS   Rattus sp.
OS   Escherichia coli.
OS   Chimeric.
XX
PN   WO200064247-A1.
XX
PD   02-NOV-2000.
XX
PF   20-APR-2000; 2000WO-CA000430.
XX
PR   23-APR-1999;   99US-0130508P.
XX

```
Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy  421 IVNEARIPACSL 432
          ||||||||||||
Db  421 IVNEARIPACSL 432
```

RESULT 12
AAB36259
ID   AAB36259 standard; protein; 432 AA.
XX
AC   AAB36259;
XX
DT   12-SEP-2003   (revised)
DT   20-FEB-2001   (first entry)
XX
DE   R15/APPA plasmid translated sequence.
XX
KW   Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW   environmental polution; pig.
XX
OS   Rattus sp.
OS   Escherichia coli.
OS   Chimeric.
XX
PN   WO200064247-A1.
XX
PD   02-NOV-2000.
XX
PF   20-APR-2000; 2000WO-CA000430.
XX
PR   23-APR-1999;   99US-0130508P.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   Forsberg CW, Golovan S, Phillips JP;
XX
DR   WPI; 2000-687245/67.
DR   N-PSDB; AAC68296.
XX
PT   Transgenic non-human animal for gastrointestinal tract specific
PT   expression of a protein, preferably phytase, comprises a nucleic acid
PT   sequence including a heterologous transgene construct encoding the
PT   protein.
XX
PS   Disclosure; Fig 19; 152pp; English.
XX

```
PA   (UYGU-) UNIV GUELPH.
XX
PI   Forsberg CW, Golovan S, Phillips JP;
XX
DR   WPI; 2000-687245/67.
DR   N-PSDB; AAC68295.
XX
PT   Transgenic non-human animal for gastrointestinal tract specific
PT   expression of a protein, preferably phytase, comprises a nucleic acid
PT   sequence including a heterologous transgene construct encoding the
PT   protein.
XX
PS   Disclosure; Fig 18; 152pp; English.
XX
CC   The present invention provides transgenic animals which produce desired
CC   proteins, in this case pigs which expresses phytase in the salivary
CC   gland. Low phytase production levels result in phytate in the diet being
CC   excreted and causing phosphorus contamination in water, as well as
CC   reducing the growth of animals. The invention provides a number of
CC   transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC   on 12-SEP-2003 to standardise OS field)
XX
SQ   Sequence 432 AA;

Query Match             97.1%;  Score 2235;  DB 3;  Length 432;
Best Local Similarity   99.5%;  Pred. No. 1.6e-216;
Matches  430;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy          1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy         61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy        121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADPTGH 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy        181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy        241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy        301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy        361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy        421 IVNEARIPACSL 432
              ||||||||||||
Db        421 IVNEARIPACSL 432

RESULT 14
AAB36263
ID   AAB36263 standard; protein; 432 AA.
XX
AC   AAB36263;
XX
DT   12-SEP-2003   (revised)
DT   20-FEB-2001   (first entry)
XX
DE   Lama2/APPA plasmid translated sequence.
XX
KW   Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW   environmental pollution; pig.
XX
OS   Mus musculus.
OS   Escherichia coli.
OS   Chimeric.
XX
PN   WO20006247-A1.
XX
PD   02-NOV-2000.
XX
PF   20-APR-2000; 2000WO-CA000430.
XX
PR   23-APR-1999;  99US-0130508P.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   Forsberg CW, Golovan S, Phillips JP;
XX
DR   WPI; 2000-687245/67.
DR   N-PSDB; AAC68300.
XX
PT   Transgenic non-human animal for gastrointestinal tract specific
PT   expression of a protein, preferably phytase, comprises a nucleic acid
PT   sequence including a heterologous transgene construct encoding the
PT   protein.
XX
PS   Disclosure; Fig 23; 152pp; English.
XX
CC   The present invention provides transgenic animals which produce desired
CC   protein, in this case pigs which expresses phytase in the salivary
CC   gland. Low phytase production levels result in phytate in the diet being
CC   excreted and causing phosphorus contamination in water, as well as
CC   reducing the growth of animals. The invention provides a number of
CC   transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC   on 12-SEP-2003 to standardise OS field)
XX
SQ   Sequence 432 AA;

Query Match             97.1%;  Score 2235;  DB 3;  Length 432;
Best Local Similarity   99.5%;  Pred. No. 1.6e-216;
Matches  430;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy          1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy         61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy        121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADPTGH 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy        181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy        241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy        301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy        361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy        421 IVNEARIPACSL 432
              ||||||||||||
Db        421 IVNEARIPACSL 432
```

```
RESULT 15
AAB36260
ID    AAB36260  standard; protein; 432 AA.
XX
AC    AAB36260;
XX
DT    12-SEP-2003  (revised)
DT    20-FEB-2001  (first entry)
XX
DE    R15/APPA plasmid translated sequence.
XX
KW    Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW    environmental pollution; pig.
XX
OS    Rattus sp.
OS    Escherichia coli.
OS    Chimeric.
XX
PN    WO20006424-A1.
XX
PD    02-NOV-2000.
XX
PF    20-APR-2000; 2000WO-CA000430.
XX
PR    23-APR-1999;  99US-0130508P.
XX
PA    (UYGU-) UNIV GUELPH.
XX
PI    Forsberg CW,  Golovan S,  Phillips JP;
XX
DR    WPI; 2000-687245/67.
DR    N-PSDB; AAC68297.
XX
PT    Transgenic non-human animal for gastrointestinal tract specific
PT    expression of a protein, preferably phytase, comprises a nucleic acid
PT    sequence including a heterologous transgene construct encoding the
PT    protein.
XX
PS    Disclosure; Fig 20; 152pp; English.
XX
CC    The present invention provides transgenic animals which produce desired
CC    proteins, in this case pigs which expresses phytase in the salivary
CC    gland. Low phytase production levels result in phytate in the diet being
CC    excreted and causing phosphorus contamination in water, as well as
CC    reducing the growth of animals. The invention provides a number of
CC    transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC    on 12-SEP-2003 to standardise OS field)
XX
SQ    Sequence 432 AA;

Query Match        97.1%;  Score 2235;  DB 3;  Length 432;
Best Local Similarity  99.5%;  Pred. No. 1.6e-216;
Matches  430;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy     1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP 60

Qy    61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy   121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy   241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGGQPDNTPPGGELVFERW 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGGQPDNTPPGGELVFERW 360

Qy   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy   421 IVNEARIPACSL 432
           ||||||||||||
Db   421 IVNEARIPACSL 432


Search completed: June 13, 2006, 10:22:03
Job time : 141.266 secs
```

THIS PAGE LEFT BLANK

```
                    GenCore version 5.1.9
               Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on:     June 13, 2006, 10:17:34 ; Search time 192.752 Seconds
                                       (without alignments)
                                       2111.554 Million cell updates/sec

Title:          US-10-601-319-2
Perfect score:  2302
Sequence:       1 MKAILIPFLSLLIPLTPQSA.........IVNEARIPACSLRSHHHHHH 440

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       2849598 seqs, 925001592 residues

Total number of hits satisfying chosen parameters:     2849598

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    UniProt 7.2:*
              1: uniprot_sprot:*
              2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	2235	97.1	432	1	PPA_ECOLI	P07102 escherichia
2	2235	97.1	432	2	Q3Z3B1_SHISS	Q3Z3B1 shigella so
3	2231	96.9	432	2	Q8GN88_ECOLI	Q8GN88 escherichia
4	2229	96.8	432	2	Q8RKD6_ECOLI	Q8RKD6 escherichia
5	2229	96.8	432	2	Q8RKD7_ECOLI	Q8RKD7 escherichia
6	2229	96.8	432	2	Q8RKD8_ECOLI	Q8RKD8 escherichia
7	2229	96.8	432	2	Q8RKE0_ECOLI	Q8RKE0 escherichia
8	2228	96.8	432	2	Q31YP3_SHIBS	Q31YP3 shigella bo
9	2228	96.8	432	2	Q8RKD9_ECOLI	Q8RKD9 escherichia
10	2225	96.7	432	2	Q8RKD5_ECOLI	Q8RKD5 escherichia
11	2222	96.5	432	2	Q83RW2_SHIFL	Q83RW2 shigella fl
12	2221	96.5	432	2	Q6RKO8_ECOLI	Q6RK08 escherichia
13	2218	96.4	432	2	Q7UD08_SHIFL	Q7UD08 shigella fl
14	2213	96.1	446	2	Q8CW75_ECOL6	Q8CW75 escherichia
15	2210	96.0	434	2	Q7AFW3_ECO57	Q7AFW3 escherichia
16	2210	96.0	444	2	Q8XC29_ECO57	Q8XC29 escherichia
17	2191	95.2	434	2	Q32HS7_SHIDS	Q32HS7 shigella dy
18	1374.5	59.7	433	2	Q676V7_CITFR	Q676V7 citrobacter
19	1361.5	59.1	433	2	Q2VY22_CITBR	Q2VY22 citrobacter
20	1077	46.8	444	2	Q6U677_9ENTR	Q6U677 obesumbacte
21	1054.5	45.8	444	2	Q6TAQ8_9ENTR	Q6TAQ8 obesumbacte
22	943.5	41.0	441	2	Q669R3_YERPS	Q669R3 yersinia ps
23	943.5	41.0	441	2	Q82FP6_YERPE	Q82FP6 yersinia pe
24	545	23.7	413	2	Q8XBZ6_ECO57	Q8XBZ6 escherichia
25	541	23.5	413	2	Q323B9_SHISS	Q323B9 shigella so
26	541	23.5	417	2	AGP_PRORE	Q52309 providencia
27	539	23.4	413	1	AGP_ECOLI	P19926 escherichia
28	538	23.4	413	2	Q83RV6_SHIFL	Q83RV6 shigella fl
29	533	23.2	413	2	Q32HQ8_SHIDS	Q32HQ8 shigella dy
30	530	23.0	413	2	Q5PG89_SALPA	Q5PG89 salmonella
31	529	23.0	413	1	AGP_SALTY	O33921 salmonella
32	529	23.0	413	2	Q57QN8_SALCH	Q57QN8 salmonella
33	526	22.8	414	2	Q9AAQ4_CAUCR	Q9AAQ4 caulobacter
34	524	22.8	413	2	Q827P1_SALTI	Q827P1 salmonella
35	523	22.7	413	2	Q6EV19_ENTCL	Q6EV19 enterobacte
36	505	21.9	392	2	Q8PP76_XANAC	Q8PP76 xanthomonas
37	505	21.9	443	2	Q4UR06_XANC8	Q4UR06 xanthomonas
38	505	21.9	443	2	Q8P330_XANCP	Q8P330 xanthomonas
39	483	21.0	515	2	Q3BXB8_XANC5	Q3BXB8 xanthomonas
40	480.5	20.9	435	2	Q8PF53_XANAC	Q8PF53 xanthomonas
41	476.5	20.7	433	2	Q3BMQ7_XANC5	Q3BMQ7 xanthomonas
42	474	20.6	532	2	Q5GW75_XANOR	Q5GW75 xanthomonas
43	467.5	20.3	435	2	Q6CZF4_ERWCT	Q6CZF4 erwinia car
44	439	19.1	427	2	Q4ZMU2_PSEU2	Q4ZMU2 pseudomonas
45	433.5	18.8	428	2	Q8GD20_PSSBX	Q8GD20 pseudomonas

ALIGNMENTS

```
RESULT 1
PPA_ECOLI
ID   PPA_ECOLI    STANDARD;    PRT;    432 AA.
AC   P07102;
DT   01-APR-1988, integrated into UniProtKB/Swiss-Prot.
DT   01-AUG-1991, sequence version 2.
DT   07-MAR-2006, entry version 63.
DE   Periplasmic appA protein precursor [Includes: Phosphoanhydride
DE   phosphohydrolase (EC 3.1.3.2) (pH 2.5 acid phosphatase) (AP); 4-
DE   phytase (EC 3.1.3.26)].
GN   Name=appA; OrderedLocusNames=b0980;
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND PROTEIN SEQUENCE OF 23-29.
RC   STRAIN=K12;
RX   MEDLINE=90368616; PubMed=2168385;
RA   Dassa J., Marck C., Boquet P.L.;
RT   "The complete nucleotide sequence of the Escherichia coli gene appA
RT   reveals significant homology between pH 2.5 acid phosphatase and
RT   glucose-1-phosphatase.";
RL   J. Bacteriol. 172:5497-5500(1990).
RN   [2]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [3]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=K12 / W3110;
RX   MEDLINE=97061202; PubMed=8905232; DOI=10.1093/dnares/3.3.137;
RA   Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA   Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA   Kimura S., Kitagawa M., Makino K., Miki T., Mizobuchi K.,
RA   Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA   Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA   Yano M., Horiuchi T.;
RT   "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 12.7-28.0 min region on the linkage map.";
RL   DNA Res. 3:137-155(1996).
RN   [4]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 1-112.
RX   MEDLINE=87271766; PubMed=3038201; DOI=10.1016/0300-9084(87)90045-9;
RA   Touati E., Danchin A.;
RT   "The structure of the promoter and amino terminal region of the pH 2.5
RT   acid phosphatase structural gene (appA) of E. coli: a negative control
RT   of transcription mediated by cyclic AMP.";
```

```
RL   Biochimie 69:215-221(1987).
RN   [5]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 1-17.
RC   STRAIN=K12;
RX   MEDLINE=92049231; PubMed=1658595; DOI=10.1007/BF00267454;
RA   Dassa J., Fsihi H., Marck C., Dion M., Kieffer-Bontemps M.,
RA   Boquet P.L.;
RT   "A new oxygen-regulated operon in Escherichia coli comprises the genes
RT   for a putative third cytochrome oxidase and for pH 2.5 acid
RT   phosphatase (appA).";
RL   Mol. Gen. Genet. 229:341-352(1991).
RN   [6]
RP   CHARACTERIZATION, AND PROTEIN SEQUENCE OF 23-34.
RA   Greiner R., Jany K.-D.;
RT   "Characterization of a phytase from Escherichia coli.";
RL   Biol. Chem. Hoppe-Seyler 372:664-665(1991).
RN   [7]
RP   CHARACTERIZATION, AND PROTEIN SEQUENCE OF 23-35.
RX   MEDLINE=93266556; PubMed=8387749; DOI=10.1006/abbi.1993.1261;
RA   Greiner R., Konietzny U., Jany K.-D.;
RT   "Purification and characterization of two phytases from Escherichia
RT   coli.";
RL   Arch. Biochem. Biophys. 303:107-113(1993).
RN   [8]
RP   CHARACTERIZATION.
RX   MEDLINE=20161462; PubMed=10696472; DOI=10.1139/cjm-46-1-59;
RA   Golovan S., Wang G., Zhang J., Forsberg C.W.;
RT   "Characterization and overproduction of the Escherichia coli appA
RT   encoded bifunctional enzyme that exhibits both phytase and acid
RT   phosphatase activities.";
RL   Can. J. Microbiol. 46:59-71(2000).
RN   [9]
RP   MUTAGENESIS.
RX   MEDLINE=93054596; PubMed=1429631;
RA   Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
RA   van Etten R.L.;
RT   "Overexpression, site-directed mutagenesis, and mechanism of
RT   Escherichia coli acid phosphatase.";
RL   J. Biol. Chem. 267:22830-22836(1992).
RN   [10]
RP   X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS).
RX   MEDLINE=20122624; PubMed=10655611; DOI=10.1038/72371;
RA   Lim D., Golovan S., Forsberg C.W., Jia Z.;
RT   "Crystal structures of Escherichia coli phytase and its complex with
RT   phytate.";
RL   Nat. Struct. Biol. 7:108-113(2000).
CC   -!- CATALYTIC ACTIVITY: A phosphate monoester + H(2)O = an alcohol +
CC       phosphate.
CC   -!- CATALYTIC ACTIVITY: Myo-inositol hexakisphosphate + H(2)O = 1D-
CC       myo-inositol 1,2,3,4,5-pentakisphosphate + phosphate.
CC   -!- SUBUNIT: Monomer.
CC   -!- SUBCELLULAR LOCATION: Periplasmic.
CC   -!- INDUCTION: In addition to cAMP-mediated control, this enzyme is
CC       induced when bacterial cultures reach stationary phase; its
CC       synthesis is triggered by phosphate starvation or a shift from
CC       aerobic to anaerobic conditions.
CC   -!- SIMILARITY: Belongs to the histidine acid phosphatase family.
CC   -----------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   -----------------------------------------------------------------------
DR   EMBL; M58708; AAA72086.1; -; Genomic_DNA.
DR   EMBL; U00096; AAC74065.1; -; Genomic_DNA.
DR   EMBL; D90735; BAA35745.1; -; Genomic_DNA.
DR   EMBL; X05471; CAA29031.1; -; Genomic_DNA.
DR   EMBL; M58708; -; NOT_ANNOTATED_CDS; Genomic_DNA.
DR   EMBL; S63811; AAB20286.1; -; Genomic_DNA.
DR   PIR; B36733; B36733.
DR   PDB; 1DKL; X-ray; A/B=23-432.
DR   PDB; 1DKM; X-ray; A=23-432.
DR   PDB; 1DKN; X-ray; A=23-432.
DR   PDB; 1DKO; X-ray; A=23-432.
DR   PDB; 1DKP; X-ray; A=23-432.
DR   PDB; 1DKQ; X-ray; A=23-432.
DR   SWISS-2DPAGE; P07102; COLI.
DR   GenomeReviews; U00096_GR; b0980.
DR   EchoBASE; EB0047; -.
DR   EcoGene; EG10049; appA.
DR   BioCyc; EcoCyc:APPA-MONOMER; -.
DR   InterPro; IPR000560; HisAc_phsphtse.
DR   Pfam; PF00328; Acid_phosphat_A; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW   3D-structure; Complete proteome; Direct protein sequencing; Hydrolase;
KW   Multifunctional enzyme; Periplasmic; Signal.
FT   SIGNAL        1     22
FT   CHAIN        23    432       Periplasmic appA protein.
FT                               /FTId=PRO_0000023947.
FT   ACT_SITE     39     39       Nucleophile.
FT   ACT_SITE    326    326       Proton donor.
FT   DISULFID     99    130
FT   DISULFID    155    430
FT   DISULFID    200    210
FT   DISULFID    404    413
FT   CONFLICT     51     66       MQDVTPDAWPTWPVKL -> NAGCHPRRMANLAGKT (in
FT                               Ref. 3).
FT   CONFLICT     75     76       EL -> DV (in Ref. 4).
FT   CONFLICT    112    112       D -> S (in Ref. 4).
FT   STRAND       27     38
FT   STRAND       41     42
FT   STRAND       45     45
FT   HELIX        49     53
FT   TURN         54     54
FT   STRAND       56     57
FT   STRAND       63     64
FT   TURN         66     67
FT   STRAND       69     69
FT   HELIX        71     90
FT   TURN         91     92
FT   STRAND       93     94
FT   STRAND       96     98
FT   TURN        102    104
FT   STRAND      105    109
FT   STRAND      111    112
FT   HELIX       113    126
FT   STRAND      127    127
FT   TURN        128    129
FT   STRAND      134    135
FT   STRAND      138    138
FT   TURN        140    141
FT   STRAND      142    142
FT   HELIX       145    147
FT   TURN        149    153
FT   STRAND      154    155
FT   HELIX       159    169
FT   TURN        170    172
FT   HELIX       174    179
FT   TURN        180    181
FT   HELIX       182    192
FT   TURN        193    193
FT   HELIX       194    196
FT   STRAND      197    197
FT   HELIX       198    201
FT   TURN        202    202
FT   STRAND      204    204
FT   TURN        206    207
FT   STRAND      208    208
FT   HELIX       212    215
FT   STRAND      216    216
FT   STRAND      220    222
FT   STRAND      224    225
FT   STRAND      227    229
FT   TURN        231    231
FT   HELIX       232    249
FT   TURN        250    250
FT   STRAND      252    253
```

```
FT   HELIX       254    257
FT   TURN        258    258

Query Match             99.5%;
Best Local Similarity   99.5%;  Score 2235; DB 1; Length 432;
  Matches  430;  Conservative    0;  Mismatches    2;  Indels    2;  Gaps    0;

Qy    1 MKAILIPFLSLLIPLTPQSAPAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAPAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP   60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120

Qy  121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180

Qy  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA  300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPGEVKLTLAGCEERNAQGMCSLAGFTQ  420

Qy  421 IVNEARIPACSL  432
        ||||||||||||
Db  421 IVNEARIPACSL  432
```

```
RESULT 2
Q3Z3E1_SHISS
ID   Q3Z3E1_SHISS     PRELIMINARY;     PRT;    432 AA.
AC   Q3Z3E1;
DT   27-SEP-2005, integrated into UniProtKB/TrEMBL.
DT   27-SEP-2005, sequence version 1.
DT   07-FEB-2006, entry version 5.
DE   Phosphoanhydride phosphorylase.
GN   Name=appA; OrderedLocusNames=SSO_0987; ORFNames=SSO_0987;
OS   Shigella sonnei (strain Ss046).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=300269;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   PubMed=16275786; DOI=10.1093/nar/gki954;
RA   Yang F., Yang J., Zhang X., Chen L., Jiang Y., Yan Y., Tang X.,
RA   Wang J., Xiong Z., Dong J., Xue Y., Zhu Y., Xu X., Sun L., Chen S.,
RA   Nie H., Peng J., Xu J., Wang Y., Yuan Z., Wen Y., Yao Z., Shen Y.,
RA   Qiang B., Hou Y., Yu J., Jin Q.;
RT   "Genome dynamics and diversity of Shigella species, the etiologic
RT   agents of bacillary dysentery.";
RL   Nucleic Acids Res. 33:6445-6458(2005).
CC   -----------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   -----------------------------------------------------------------------
DR   EMBL; CP000038; AAZ87721.1; -; Genomic_DNA.
DR   GO; GO:0003993; F:acid phosphatase activity; IEA.
DR   InterPro; IPR000560; HisAc_phsphtse.
DR   Pfam; PF00328; Acid_phosphat_A; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
```

```
KW   Complete proteome.
SQ   SEQUENCE   432 AA;  47057 MW;  6510C6C579177F11 CRC64;

Query Match             97.1%;
Best Local Similarity   99.5%;  Pred. No. 7.6e-163;  DB 2;  Length 432;
  Matches  430;  Conservative    0;  Mismatches    2;  Indels    2;  Gaps    0;

Qy    1 MKAILIPFLSLLIPLTPQSAPAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAPAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP   60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120

Qy  121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180

Qy  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA  300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPGEVKLTLAGCEERNAQGMCSLAGFTQ  420

Qy  421 IVNEARIPACSL  432
        ||||||||||||
Db  421 IVNEARIPACSL  432
```

```
RESULT 3
Q8GN88_ECOLI
ID   Q8GN88_ECOLI     PRELIMINARY;     PRT;    432 AA.
AC   Q8GN88;
DT   01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT   01-MAR-2003, sequence version 1.
DT   07-FEB-2006, entry version 11.
DE   AppA.
GN   Name=appA;
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Zhang Z., He J., Yao B., Zhou Y., Chen Y., Yi Y.;
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Chen Y., Zhu J., Zhang Z., He J.;
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
CC   -----------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   -----------------------------------------------------------------------
DR   EMBL; AF537219; AAN28334.1; -; Genomic_DNA.
DR   HSSP; P07102; 1DKM.
DR   SMR; Q8GN88; 23-432.
DR   GO; GO:0003993; F:acid phosphatase activity; IEA.
DR   InterPro; IPR000560; HisAc_phsphtse.
DR   Pfam; PF00328; Acid_phosphat_A; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
```

DR   HSSP; P07102; 1DKM.
DR   SMR; Q8RKD6; 23-432.
DR   GO; GO:0003993; F:acid phosphatase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   InterPro; IPR000560; HisAc_phsphtse.
DR   Pfam; PF00328; Acid_phosphat_A; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW   Hydrolase.
SQ   SEQUENCE   432 AA; 46972 MW; AF86C41EA6193AC5 CRC64;

Query Match   96.8%; Score 2229; DB 2; Length 432;
Best Local Similarity   99.3%; Pred. No. 2.2e-162;
Matches 429; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

Qy   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDEATRKTGE 120

Qy   121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy   241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy   421 IVNEARIPACSL 432
          ||||||||||||
Db   421 IVNEARIPACSL 432

RESULT 5
Q8RKD7_ECOLI
ID   Q8RKD7_ECOLI   PRELIMINARY;   PRT;   432 AA.
AC   Q8RKD7;
DT   01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT   01-JUN-2002, sequence version 1.
DT   07-FEB-2006, entry version 13.
DE   Periplasmic phosphoanhydride phosphohydrolase.
GN   Name=appA;
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=90368616; PubMed=2168385;
RA   Dassa J., Marck C., Boquet P.L.;
RT   "The complete nucleotide sequence of the Escherichia coli gene appA
RT   reveals significant homology between pH 2.5 acid phosphatase and
RT   glucose-1-phosphatase.";
RL   J. Bacteriol. 172:5497-5500(1990).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=93054596; PubMed=1429631;
RA   Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,

---

SQ   SEQUENCE   432 AA; 47056 MW; 5B355D76B7377737 CRC64;

Query Match   96.9%; Score 2231; DB 2; Length 432;
Best Local Similarity   99.3%; Pred. No. 1.5e-162;
Matches 429; Conservative   1; Mismatches   2; Indels   0; Gaps   0;

Qy   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy   121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy   241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
          :||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 KIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPCGGELVFERW 360

Qy   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy   421 IVNEARIPACSL 432
          ||||||||||||
Db   421 IVNEARIPACSL 432

RESULT 4
Q8RKD6_ECOLI
ID   Q8RKD6_ECOLI   PRELIMINARY;   PRT;   432 AA.
AC   Q8RKD6;
DT   01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT   01-JUN-2002, sequence version 1.
DT   07-FEB-2006, entry version 13.
DE   Periplasmic phosphoanhydride phosphohydrolase.
GN   Name=appA;
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=93054596; PubMed=1429631;
RA   Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
RA   van Etten R.L.;
RT   "Overexpression, site-directed mutagenesis, and mechanism of
RT   Escherichia coli acid phosphatase.";
RL   J. Biol. Chem. 267:22830-22836(1992).
CC   ---
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---
DR   EMBL; L03374; AAA00006.1; -; Genomic_DNA.

Left column

RA van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase.";
RL J. Biol. Chem. 267:22830-22836(1992).

DR EMBL; L03373; AAA00005.1; -; Genomic_DNA.
DR HSSP; P07102; 1DKM.
DR SMR; Q8RKD7; 23-432.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000560; HisAc_phsphtse.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase.
SQ SEQUENCE 432 AA; 46972 MW; 755D5E4B1AD9916A6 CRC64;

Query Match   96.8%; Score 2229; DB 2; Length 432;
Best Local Similarity 99.3%; Pred. No. 2.2e-162;
Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGILLAKKGCPQSGQVAIIADVDERTRKTGE 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGILLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKYSADNVSLTGAVSLASMLT 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKYSADNVSLTGAVSLASMLT 240

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy  421 IVNEARIPACSL 432
        ||||||||||||
Db  421 IVNEARIPACSL 432
```

RESULT 6
Q8RKD8_ECOLI
ID   Q8RKD8_ECOLI PRELIMINARY; PRT; 432 AA.
AC   Q8RKD8;
DT   01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT   01-JUN-2002, sequence version 1.
DT   07-FEB-2006, entry version 13.
DE   Periplasmic phosphoanhydride phosphohydrolase.
GN   Name=appA;
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=90368616; PubMed=2168385;

Right column

RA Dassa J., Marck C., Boquet P.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase.";
RL J. Bacteriol. 172:5497-5500(1990).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
RA van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase.";
RL J. Biol. Chem. 267:22830-22836(1992).

DR EMBL; L03372; AAA00004.1; -; Genomic_DNA.
DR HSSP; P07102; 1DKM.
DR SMR; Q8RKD8; 23-432.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000560; HisAc_phsphtse.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase.
SQ SEQUENCE 432 AA; 46972 MW; 9A85536B57FCCFB5 CRC64;

Query Match   96.8%; Score 2229; DB 2; Length 432;
Best Local Similarity 99.3%; Pred. No. 2.2e-162;
Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGILLAKKGCPQSGQVAIIADVDERTRKTGE 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGILLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKYSADNVSLTGAVSLASMLT 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKYSADNVSLTGAVSLASMLT 240

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy  421 IVNEARIPACSL 432
        ||||||||||||
Db  421 IVNEARIPACSL 432
```

RESULT 7
Q8RKE0_ECOLI
ID   Q8RKE0_ECOLI PRELIMINARY; PRT; 432 AA.
AC   Q8RKE0;
DT   01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT   01-JUN-2002, sequence version 1.
DT   07-FEB-2006, entry version 13.
DE   Periplasmic phosphoanhydride phosphohydrolase.

```
GN   Name=appA;
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=93054596; PubMed=1429631;
RA   Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
RA   van Etten R.L.;
RT   "Overexpression, site-directed mutagenesis, and mechanism of
RT   Escherichia coli acid phosphatase.";
RL   J. Biol. Chem. 267:22830-22836(1992).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=93054596; PubMed=2168385;
RA   Dassa J., Marck C., Boquet P.L.;
RT   "The complete nucleotide sequence of the Escherichia coli gene appA
RT   reveals significant homology between pH 2.5 acid phosphatase and
RT   glucose-1-phosphatase.";
RL   J. Bacteriol. 172:5497-5500(1990).
```
```
DR   EMBL; L03370; AAA00002.1; -; Genomic_DNA.
DR   SMR; Q8RKE0; 23-432.
DR   GO; GO:0003993; F:acid phosphatase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   InterPro; IPR000560; HiBAC_phsphtse.
DR   Pfam; PF00328; Acid_phosphat_A; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW   Hydrolase.
SQ   SEQUENCE   432 AA;  46972 MW;  5BBD632D4682EABF CRC64;

Query Match     96.8%;   Score 2229;  DB 2;  Length 432;
Best Local Similarity   99.3%;   Pred. No. 2.2e-162;
Matches  429;  Conservative  0;  Mismatches  3;  Indels  0;  Gaps  0;

Qy      1  MKAILIPFLSLLIPLTPQSAFAQSEEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  MKAILIPFLSLLIPLTPQSAFAQSEEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60

Qy     61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120

Qy    121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180

Qy    181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240

Qy    241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA  300

Qy    301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW  360

Qy    361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420

Qy    421  IVNEARIPACSL  432
           ||||||||||||
Db    421  IVNEARIPACSL  432
```

```
RESULT 8

ID   Q31YP3_SHIBS            PRELIMINARY;      PRT;   432 AA.
AC   Q31YP3;
DT   06-DEC-2005, integrated into UniProtKB/TrEMBL.
DT   06-DEC-2005, sequence version 1.
DT   07-FEB-2006, entry version 4.
DE   Phosphoanhydride phosphorylase.
GN   Name=appA; OrderedLocusNames=SBO_2250;
OS   Shigella boydii serotype 4 (strain Sb227).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=300268;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   PubMed=16275786; DOI=10.1093/nar/gki954;
RA   Yang F., Yang J., Zhang X., Chen L., Jiang Y., Yan Y., Tang X.,
RA   Wang J., Xiong Z., Dong J., Xue Y., Zhu Y., Xu X., Sun L., Chen S.,
RA   Nie H., Peng J., Xu J., Wang Y., Yuan Z., Wen Y., Yao Z., Shen Y.,
RA   Qiang B., Hou Y., Yu J., Jin Q.;
RT   "Genome dynamics and diversity of Shigella species, the etiologic
RT   agents of bacillary dysentery.";
RL   Nucleic Acids Res. 33:6445-6458(2005).
```
```
DR   EMBL; CP000036; ABB66815.1; -; Genomic_DNA.
DR   GO; GO:0003993; F:acid phosphatase activity; IEA.
KW   Complete proteome.
SQ   SEQUENCE   432 AA;  47064 MW;  8569022BA28D9C73 CRC64;

Query Match     96.8%;   Score 2228;  DB 2;  Length 432;
Best Local Similarity   99.3%;   Pred. No. 2.6e-162;
Matches  429;  Conservative  0;  Mismatches  3;  Indels  0;  Gaps  0;

Qy      1  MKAILIPFLSLLIPLTPQSAFAQSEEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  MKAILIPFLSLLIPLTPQSAFAQSEEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60

Qy     61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120

Qy    121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180

Qy    181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240

Qy    241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300

Qy    301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW  360

Qy    361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420

Qy    421  IVNEARIPACSL  432
           ||||||||||||
Db    421  IVNEARIPACSL  432
```

```
RESULT 9
ID   Q8RKD9_ECOLI
AC   Q8RKD9;
```

DT   01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT   01-JUN-2002, sequence version 1.
DT   07-FEB-2006, entry version 13.
DE   Periplasmic phosphoanhydride phosphohydrolase.
GN   Name=appA;
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=90368616; PubMed=2168385;
RA   Dassa J., Marck C., Boquet P.L.;
RT   "The complete nucleotide sequence of the Escherichia coli gene appA
RT   reveals significant homology between pH 2.5 acid phosphatase and
RT   glucose-1-phosphatase.";
RL   J. Bacteriol. 172:5497-5500(1990).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RP   MEDLINE=93054596; PubMed=1429631;
RA   Ostanin K., Harms E.H., Stevis P.B., Kuciel R., Zhou M.-M.,
RA   van Etten R.L.;
RT   "Overexpression, site-directed mutagenesis, and mechanism of
RT   Escherichia coli acid phosphatase.";
RL   J. Biol. Chem. 267:22830-22836(1992).
CC   ---
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---
DR   EMBL; L03371; AAA00003.1; -; Genomic_DNA.
DR   HSSP; P07102; 1DKQ.
SMR;   Q8RKD9; 23-432.
GO;   GO:0003993; F:acid phosphatase activity; IEA.
GO;   GO:0016787; F:hydrolase activity; IEA.
InterPro; IPR000560; HisAc_phsphtse.
Pfam; PF00328; Acid_phosphat_A; 1.
PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW   Hydrolase.
SQ   SEQUENCE   432 AA; 47034 MW; 9F29B9DF9C368175 CRC64;

Query Match   96.8%; Score 2228; DB 2; Length 432;
Best Local Similarity   99.3%;
Matches 429; Conservative   1; Mismatches   2; Indels   0; Gaps   0;

Qy   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy   121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKYSADNVSLTGAVSLASMLT 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKYSADNVSLTGAVSLASMLT 240

Qy   241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360

Qy   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy   421 IVNEARIPACSL 432
             ||||||||||||
Db   421 IVNEARIPACSL 432

Db   421 IVNEARIPACSL 432

RESULT 10
ID   Q8RKD5_ECOLI   PRELIMINARY;   PRT;   432 AA.
AC   Q8RKD5;
DT   01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT   01-JUN-2002, sequence version 1.
DT   07-FEB-2006, entry version 13.
DE   Periplasmic phosphoanhydride phosphohydrolase.
GN   Name=appA;
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=90368616; PubMed=2168385;
RA   Dassa J., Marck C., Boquet P.L.;
RT   "The complete nucleotide sequence of the Escherichia coli gene appA
RT   reveals significant homology between pH 2.5 acid phosphatase and
RT   glucose-1-phosphatase.";
RL   J. Bacteriol. 172:5497-5500(1990).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RP   MEDLINE=93054596; PubMed=1429631;
RA   Ostanin K., Harms E.H., Stevis P.B., Kuciel R., Zhou M.-M.,
RA   van Etten R.L.;
RT   "Overexpression, site-directed mutagenesis, and mechanism of
RT   Escherichia coli acid phosphatase.";
RL   J. Biol. Chem. 267:22830-22836(1992).
CC   ---
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---
DR   EMBL; L03375; AAA00007.1; -; Genomic_DNA.
DR   HSSP; P07102; 1DKW.
SMR;   Q8RKD5; 23-432.
GO;   GO:0003993; F:acid phosphatase activity; IEA.
GO;   GO:0016787; F:hydrolase activity; IEA.
InterPro; IPR000560; HisAc_phsphtse.
Pfam; PF00328; Acid_phosphat_A; 1.
PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
KW   Hydrolase.
SQ   SEQUENCE   432 AA; 46991 MW; 951F393EA9A1A47C CRC64;

Query Match   96.7%; Score 2225; DB 2; Length 432;
Best Local Similarity   99.3%;
Matches 429; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

Qy   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy   121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKYSADNVSLTGAVSLASMLT 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKYSADNVSLTGAVSLASMLT 240

Qy   241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360

```
Db    301  LTPHPPQKQAYGVTLPTSVLFIAGADTNLANIGGALELNWTLPGQPDNTPPGGELVFERW 360
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    361  RRLSDNSQWIQVSLVFQCTLQQMRDKTPLSLNTTPGEVKLTILAGCEERNAQGMCSLAGFTQ 420
Db    361  RRLSDNSQWIQVSLVFQCTLQQMRDKTPLSLNTTPGEVKLTILAGCEERNAQGMCSLAGFTQ 420
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    421  IVNEARIPACSL 432
               ||||||||||||
Db    421  IVNEARIPACSL 432
```

RESULT 11

```
ID    Q83RW2_SHIFL   PRELIMINARY;  PRT;  432 AA.
AC    Q83RW2;
DT    01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT    05-JUL-2005, sequence version 4.
DT    07-FEB-2006, entry version 15.
DE    Phosphoanhydride phosphorylase; pH 2.5 acid phosphatase.
GN    Name=appA; OrderedLocusNames=SF0982;
OS    Shigella flexneri.
OC    Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC    Enterobacteriaceae; Shigella.
OX    NCBI_TaxID=623;
RN    [1]
RP    NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC    STRAIN=301 / Serotype 2a;
RX    MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA    Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA    Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu D., Dong J.,
RA    Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA    Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA    Yu J.;
RT    "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT    through comparison with genomes of Escherichia coli K12 and O157.";
RL    Nucleic Acids Res. 30:4432-4441(2002).
CC    -----------------------------------------------------------------------
CC    Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC    Distributed under the Creative Commons Attribution-NoDerivs License
CC    -----------------------------------------------------------------------
EMBL; AE005674; AAN42610.2; -; Genomic_DNA.
SMR; Q83RW2; 23-432.
BioCyc; SFLE198214:AAN42610.1-MONOMER; -.
GO; GO:0003993; F:acid phosphatase activity; IEA.
InterPro; IPR000560; HiBAc_phephtse.
Pfam; PF00328; Acid_phosphat_A; 2.
PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW    Complete proteome.
SQ    SEQUENCE   432 AA;  47106 MW;  25093A53392B9C18C CRC64;
```

Query Match   96.5%; Score 2222; DB 2; Length 432;
Best Local Similarity 98.8%; Pred. No. 7.6e-162;
Matches 427; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy     1  MKAILIPFLSLLIPLPTPQSAFAQSEPELKLESSVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MKAILIPFLSLLIPLPTPQSAPAQSEPELKLESMVIVSRHGVRAPTKATQLMQDVTPDAWP 60
Qy    61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
Qy   121  AFAAGLAPDCAITVHTQADTSSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  AFAAGLAPDCAITVHTQADTSSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
Qy   181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  RQTVFRELERVLNFPQSNLCLNREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy   241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQPYLLQRTPEVARSRATPLLDLIMAA 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 12

```
ID    Q6RKO8_ECOLI   PRELIMINARY;  PRT;  432 AA.
AC    Q6RKO8;
DT    05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT    05-JUL-2004, sequence version 1.
DT    07-FEB-2006, entry version 8.
DE    Acid phosphatase/phytase 2.
GN    Name=appA2;
OS    Escherichia coli.
OC    Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC    Enterobacteriaceae; Escherichia.
OX    NCBI_TaxID=562;
RN    [1]
RP    NUCLEOTIDE SEQUENCE.
RA    Lei X.G.;
RL    Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.
RN    [2]
RP    NUCLEOTIDE SEQUENCE.
RX    MEDLINE=99194564; PubMed=10092520; DOI=10.1006/bbrc.1999.0361;
RA    Rodriguez E., Han Y., Lei X.G.;
RT    "Cloning, sequencing, and expression of an Escherichia coli acid
RT    phosphatase/phytase gene (appA2) isolated from pig colon.";
RL    Biochem. Biophys. Res. Commun. 257:117-123(1999).
CC    -----------------------------------------------------------------------
CC    Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC    Distributed under the Creative Commons Attribution-NoDerivs License
CC    -----------------------------------------------------------------------
EMBL; AY496073; AAR87658.1; -; Genomic_DNA.
SMR; Q6RKO8; 23-432.
GO; GO:0003993; F:acid phosphatase activity; IEA.
InterPro; IPR000560; HiBAc_phephtse.
Pfam; PF00328; Acid_phosphat_A; 1.
PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
FT    CHAIN    20    432    acid phosphatase/phytase 2.
SQ    SEQUENCE   432 AA;  47042 MW;  71B0E5EBD2EA2674 CRC64;
```

Query Match   96.5%; Score 2221; DB 2; Length 432;
Best Local Similarity 99.1%; Pred. No. 9.1e-162;
Matches 428; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy     1  MKAILIPFLSLLIPLPTPQSAFAQSEPELKLESSVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MKAILIPFLSLLIPLPTPQSAPAQSEPELKLESSVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
Qy    61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120
Qy   121  AFAAGLAPDCAITVHTQADTSSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  AFAAGLAPDCAITVHTQADTSSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
Qy   121  AFAAGLAPDCAITVHTQADTSSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   121  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy   181  RQTAFRELERVLNFSQLNLCLNREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  RQTAFRELERVLNFSQLNLCLNREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
```

```
Qy  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300

Qy  301  LTPHPPQQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  LTPHPPQQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360

Qy  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420

Qy  421  IVNEARIPACSL  432
         ||||||||||||
Db  421  IVNEARIPACSL  432
```

RESULT 13
```
ID   Q7UD08_SHIFL    PRELIMINARY;    PRT;    432 AA.
AC   Q7UD08;
DT   01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT   01-OCT-2003, sequence version 1.
DT   07-MAR-2006, entry version 12.
DE   Phosphoanhydride phosphorylase.
GN   Name=appA; OrderedLocusNames=S1048; ORFNames=S_1048;
OS   Shigella flexneri.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=623;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX   MEDLINE=22590274; PubMed=12704152;
RX   DOI=10.1128/IAI.71.5.2775-2786.2003;
RA   Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA   Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA   Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA   Schwartz D.C., Blattner F.R.;
RT   "Complete genome sequence and comparative genomics of Shigella
RT   flexneri serotype 2a strain 2457T.";
RL   Infect. Immun. 71:2775-2786(2003).
```
```
DR   EMBL; AE014073; AAP16495.1; -; Genomic_DNA.
DR   HSSP; P07102; 1DKL.
     SMR; Q7UD08; 23-432.
DR   GenomeReviews; AE014073_GR; S1048.
DR   GO; GO:0003993; F:acid phosphatase activity; IEA.
DR   InterPro; IPR000560; HisAc_phsphtse.
DR   Pfam; PF00328; Acid_phosphat_A; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN_1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
SQ   SEQUENCE   432 AA;  47136 MW;  20044223 9BC66DFC CRC64;

Query Match             96.4%;  Score 2218;  DB 2;  Length 432;
Best Local Similarity   98.6%;
Matches  426;  Conservative  1;  Mismatches  5;  Indels  0;  Gaps  0;

Qy    1  MKAILIPFLSLLIPLPTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60
         ||||||| ||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db    1  MKAILIPPLSLLIPLPTPQSAFAQSEPELKLESMVIVSRHGVRAPTKATQLMQDVTPDAWP   60

Qy   61  TWPVKLGWLTRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  TWPVKLGWLTRGGELIAYLGHYQRQRLVADGLLAKKGCPQSAQVAIIADVDERTRKTGE  120

Qy  121  APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180
         |||| ||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db  121  AFTAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILCRAGGSIADFTGH  180
```

```
Qy  181  RQTAFRELERVLNPPQSNLCLRREKQDESCSLTQALFSELKVSADNVSLTGAVSLASMLT  240
         |||.|||||||||.||||||| ||||||||||||||:||||||||||||||||||||||||
Db  181  RQTVFRELERVLNFPQSNLCLNRREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240

Qy  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA  300

Qy  301  LTPHPPQQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  LTPHPPQQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360

Qy  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420

Qy  421  IVNEARIPACSL  432
         ||||||||||||
Db  421  IVNEARIPACSL  432
```

RESULT 14
```
ID   Q8CW75_ECOL6    PRELIMINARY;    PRT;    446 AA.
AC   Q8CW75;
DT   01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT   01-MAR-2003, sequence version 1.
DT   21-FEB-2006, entry version 15.
DE   Periplasmic appA protein.
GN   Name=appA; ORFNames=c1121;
OS   Escherichia coli O6.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=217992;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
RX   MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA   Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA   Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA   Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA   Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT   "Extensive mosaic structure revealed by the complete genome sequence
RT   of uropathogenic Escherichia coli.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
```
```
DR   EMBL; AE014075; AAN79589.1; -; Genomic_DNA.
DR   HSSP; P07102; 1DKQ.
     SMR; Q8CW75; 33-442.
DR   BioCyc; ECOL199310:C1121-MONOMER; -.
DR   GO; GO:0003993; F:acid phosphatase activity; IEA.
DR   InterPro; IPR000560; HisAc_phsphtse.
DR   Pfam; PF00328; Acid_phosphat_A; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW   Complete proteome.
SQ   SEQUENCE   446 AA;  48606 MW;  F1308CD691DB9F49 CRC64;

Query Match             96.1%;  Score 2213;  DB 2;  Length 446;
Best Local Similarity   98.4%;
Matches  425;  Conservative  3;  Mismatches  4;  Indels  0;  Gaps  0;

Qy    1  MKAILIPFLSLLIPLPTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60
         ||||||||||||||||||.|||||||||||||||||||||||||||||||||||||||||
Db   11  MKAILIPFLSLLIPLPTPKSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   70

Qy   61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPGSGQVAIISDVDERTRKTGE  120
         |||||||||||||||||||||||||||||||||:|||||||| |||||||||||||||||
Db   71  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLTKKGCPQPGQVAIISDVDERTRKTGE  130
```

```
Qy      121 APAAGLAPDCAITVHTQADTSSPPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      131 APAAGLAPDCAITVHTQADTSSPPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 190

Qy      181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      191 RQTAFRELERVLNFPQSNLCLNREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 250

Qy      241 EIFLLQQAQGMPEPGCWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      251 EIFLLQQAQGMPEPGCWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 310

Qy      301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELANWTLPGGQPDNTPPGGELVFERW 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      311 LTPHPSQKQAYGVTLPTSVLFIAGHDTNLANLGGALELANWTLPGGQPDNTPPGGELVFERW 370

Qy      361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      371 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 430

Qy      421 IVNEARIPACSL 432
            ||||||||||:|
Db      431 IVNEARIPACSL 442
```

```
RESULT 15
ID      Q7AFW3_ECO57     PRELIMINARY;   PRT;    434 AA.
AC      Q7AFW3;
DT      05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT      05-JUL-2004, sequence version 1.
DT      07-FEB-2006, entry version 10.
DE      Phosphoanhydride phosphorylase.
GN      OrderedLocusNames=ECs1136;
OS      Escherichia coli O157:H7.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Escherichia.
OX      NCBI_TaxID=83334;
RN      [1]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX      MEDLINE=21156231; PubMed=11258796; DOI=10.1093/dnares/8.1.11;
RA      Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA      Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA      Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA      Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT      "Complete genome sequence of enterohemorrhagic Escherichia coli
RT      O157:H7 and genomic comparison with a laboratory strain K-12.";
RL      DNA Res. 8:11-22(2001).
CC      ----------------------------------------------------------------------
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      ----------------------------------------------------------------------
DR      EMBL; BA000007; BAB34559.1; -; Genomic_DNA.
DR      SMR; Q7AFW3; 25-434.
DR      GO; GO:0003993; F:acid phosphatase activity; IEA.
DR      InterPro; IPR000560; HisAc_phosphtse.
DR      Pfam; PF00328; Acid_phosphat_A; 1.
DR      PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR      PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN 1.
SQ      SEQUENCE   434 AA;  47337 MW;  F197DF7D18699C4 CRC64;
```

```
Query Match             96.0%;  Score 2210;  DB 2;  Length 434;
Best Local Similarity   98.4%;  Pred. No. 6.4e-161;
Matches   427;  Conservative   0;  Mismatches   5;  Indels   2;  Gaps   1;
```

```
Qy        1 MKAILIPFLSLLIPLTPQSAFAQS--BPELKLESVVIVSRHGVRAPTKATQLMQDVTPDA 58
            ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db        1 MKAILIPFLSLLIPLTPQSAFAQSPEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDA 60

Qy       59 WPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGGQVAIIADVDERTRKT 118
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       61 WPNWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLTKKGCPQPGGQVAIIADVDERTRKT 120
            |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Qy      119 GEAFAAGLAPDCAITVHTQADTSSPPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFT 178
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 GEAFAAGLAPDCAITVHTQADTSSPPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFT 180

Qy      179 GHRQTAFRELERVLNFPQSNICLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASM 238
            ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db      181 GHRQTAFRELERVLNFPQSNLCLNREKQDESCSLTQALPSELKVSADNVSLTGAVSLASM 240

Qy      239 LTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIM 298
            ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db      241 LTEIFLLQQAQGMPEPGCWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIM 300

Qy      299 AALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELANWTLPGGQPDNTPPGGELVFE 358
            |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 IALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELANWTLPGGQPDNTPPGGELVFE 360

Qy      359 RWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGF 418
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 RWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGF 420

Qy      419 TQIVNEARIPACSL 432
            ||||||||||||||
Db      421 TQIVNEARIPACSL 434
```

Search completed: June 13, 2006, 10:28:32
Job time : 196.752 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2006, 10:17:14 ; Search time 136.734 Seconds
                           (without alignments)
                           1444.539 Million cell updates/sec

Title: US-10-601-319-10
Perfect score: 2258
Sequence: 1 MKAILIPFLSLLIPLTPQSA.........CSLAGFTQIVNEARIPACSL 432

Scoring table: BLOSUM62
                     Gapop 10.0 , Gapext 0.5
Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   A_Geneseq_8:*
                 1: geneseqp1980s:*
                 2: geneseqp1990s:*
                 3: geneseqp2000s:*
                 4: geneseqp2001s:*
                 5: geneseqp2002s:*
                 6: geneseqp2003as:*
                 7: geneseqp2003bs:*
                 8: geneseqp2004s:*
                 9: geneseqp2005s:*
                 10: geneseqp2006s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2258	100.0	432	8	Ado50300	Ado50300 Escherich
2	2258	100.0	432	9	Aed50818	Aed50818 Escherich
3	2258	100.0	432	10	Aee75421	Aee75421 E. coli B
4	2258	100.0	436	7	Adc87743	Adc87743 Modified
5	2228	98.7	430	5	Aae15808	Aae15808 Escherich
6	2228	98.7	430	6	Ada19451	Ada19451 E. coli K
7	2182	96.6	432	3	Adc87745	Adc87745 Escherich
8	2182	96.6	432	3	Aab36257	Aab36257 Lama2/APP
9	2182	96.6	432	3	Aab36261	Aab36261 R15/APPA
10	2182	96.6	432	3	Aab36262	Aab36262 SV40/APPA
11	2182	96.6	432	3	Aab36258	Aab36258 R15/APPA
12	2182	96.6	432	3	Aab36259	Aab36259 R15/APPA
13	2182	96.6	432	3	Aab36263	Aab36263 Lama2/APP
14	2182	96.6	432	3	Aab36260	Aab36260 R15/APPA
15	2182	96.6	432	3	Aau77775	Aau77775 Phytase a
16	2182	96.6	432	5	Abp53625	Abp53625 Acid phos
17	2182	96.6	432	6	Ada19450	Ada19450 E. coli K
18	2182	96.6	432	7	Abu15427	Abu15427 Protein e
19	2182	96.6	432	8	Adc87745	Adc87745 Escherich
20	2182	96.6	432	8	Adl16129	Adl16129 Escherich
21	2182	96.6	432	9	Ado50298	Ado50298 Escherich
22	2182	96.6	432	9	Aed50820	Aed50820 Escherich
23	2182	96.6	432	10	Aee75419	Aee75419 E. coli B
24	2179	96.5	432	10	Aef15220	Aef15220 Escherich
25	2177	96.4	432	4	Aae02631	Aae02631 E. coli a
26	2177	96.4	432	8	Adl16131	Adl16131 Shigella
27	2172	96.2	440	8	Ado50304	Ado50304 Escherich
28	2171	96.1	432	4	Aae02634	Aae02634 E. coli a
29	2171	96.1	440	4	Aae02634	Aae02634 E. coli a
30	2171	96.1	440	5	Aab37892	Aab37892 Escherich
31	2171	96.1	440	6	Aae22836	Aae22836 Escherich
32	2171	96.1	440	6	Aae15806	Aae15806 Escherich
33	2171	96.1	440	6	Ada19446	Ada19446 E. coli B
34	2171	96.1	440	8	Ado50292	Ado50292 Escherich
35	2170	96.1	432	10	Aee75413	Aee75413 E. coli B
36	2169	96.1	432	4	Aae02635	Aae02635 E. coli a
37	2168.5	96.0	440	8	Ado50302	Ado50302 Kangaroo
38	2168.5	96.0	431	6	Aae37853	Aae37853 pNOV4054
39	2168.5	96.0	431	6	Aed46568	Aed46568 Nov9x phy
40	2168.5	96.0	437	6	Aae37854	Aae37854 pNOV4058
41	2168.5	96.0	437	9	Aec39116	Aec39116 Maize 27
42	2157	95.5	412	6	Aed46570	Aed46570 Nov9x phy
43	2157	95.5	412	7	Aae37851	Aae37851 Maize-opt
44	2157	95.5	432	6	Abr82310	Abr82310 Nov9X phy
45	2153	95.3	410	5	Abp51937	Abp51937 Phytase p

ALIGNMENTS

RESULT 1
Ado50300

ID   Ado50300 standard; protein; 432 AA.
XX
AC   Ado50300;
XX
DT   29-JUL-2004 (first entry)
XX
DE   Escherichia coli K12 appA phytase mutant 819PH59.
XX
KW   AppA phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW   phytate; animal feed; fish feed; dough; baking; enzyme; mutant; mutein.
OS   Escherichia coli; K12.
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FH
FT   Misc-difference   68
FT   /note= "Wild-type Trp replaced with Glu"
FT   Misc-difference   84
FT   /note= "Wild-type Gln replaced with Trp"
FT   Misc-difference   95
FT   /note= "Wild-type Ala replaced with Pro"
FT   Misc-difference   97
FT   /note= "Wild-type Lys replaced with Cys"
FT   Misc-difference   168
FT   /note= "Wild-type Ser replaced with Glu"
FT   Misc-difference   181
FT   /note= "Wild-type Arg replaced with Tyr"
FT   Misc-difference   226
FT   /note= "Wild-type Asn replaced with Cys"
FT   Misc-difference   277
FT   /note= "Wild-type Tyr replaced with Asp"
XX
PN   US20040091968-A1.
XX
PD   13-MAY-2004.
XX
PF   20-JUN-2003; 2003US-00601319.
XX
PR   13-AUG-1997; 97US-00910798.
PR   01-MAR-1999; 99US-00259214.
PR   13-APR-1999; 99US-00291931.
PR   25-MAY-1999; 99US-00318528.
PR   25-MAY-2000; 2000US-00580515.
PR   24-MAY-2001; 2001US-00866379.

```
XX
PA  (SHOR/)  SHORT J M.
PA  (KRET/)  KRETZ K.
PA  (GRAY/)  GRAY K A.
PA  (BART/)  BARTON N R.
PA  (GARR/)  GARRETT J B.
PA  (ODON/)  O'DONOGHUE E.
PA  (MATH/)  MATHER E J.
XX
PI  Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI  Mather EJ;
XX
DR  WPI; 2004-374952/35.
DR  N-PSDB; AD050299.
XX
PT  Producing phytase, involves providing nucleic acid derived from bacteria
PT  encoding polypeptide having phytase activity, and expressing nucleic acid
PT  in yeast.
XX
PS  Claim 3; SEQ ID NO 10; 74pp; English.
XX
CC  The invention relates to a method for producing phytase that involves
CC  providing a nucleic acid encoding phytase derived from a bacteria, and
CC  expressing the nucleic acid in a yeast under conditions that allow
CC  expression of the enzyme in the yeast. The invention also relates to
CC  modified phytase enzyme which has improved thermal tolerance and protease
CC  stability at low pH. The phytase enzyme can be used in foodstuffs to
CC  improve the feeding value of phytate rich ingredients, and in diet of
CC  numerous animals including mammals, fowls and fishes, commercially
CC  significant mammals such as pigs, goats, laboratory rodents, commercially
CC  significant avian species such as chicken, ducks, doves, parrot, etc.,
CC  commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC  etc; in dough making and baking, in dietary aids for animals. The method
CC  provides easy manufacture of the active ingredient loaded biocompatible
CC  composition, higher yields and loading efficiency. The phytase
CC  incorporated in to the dietary aid is safe for animals. The present
CC  sequence is Escherichia coli K12 appA phytase enzyme mutant.
XX
SQ  Sequence 432 AA;

Query Match  100.0%;  Score 2258;  DB 8;  Length 432;
Best Local Similarity  100.0%;  Pred. No. 3.2e-219;
Matches  432;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60

Qy   61  TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE  120

Qy  121  AFAAGLAPDCAITVHTQADTSSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADTGH  180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  AFAAGLAPDCAITVHTQADTSSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADTGH  180

Qy  181  YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT  240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT  240

Qy  241  EIFLLQQAQGMPEPGWGRITTDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLLDLIKTA  300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  EIFLLQQAQGMPEPGWGRITTDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLLDLIKTA  300

Qy  301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNIANLGGALELNWTLPGQPDNTPPGGELVPERW  360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNIANLGGALELNWTLPGQPDNTPPGGELVPERW  360

Qy  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420

Qy  421  IVNEARIPACSL  432
         ||||||||||||
Db  421  IVNEARIPACSL  432
```

```
RESULT 2
AED50818
ID   AED50818 standard; protein; 432 AA.
XX
AC   AED50818;
XX
DT   29-DEC-2005  (first entry)
XX
DE   Escherichia coli phytase mutant protein.
XX
KW   Transgenic animal; gene therapy; feedstuff; food; thermostable; enzyme;
KW   EC 3.1.2.8; 3-phytase; recombinant protein; mutein.
XX
OS   Escherichia coli; strain K12.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference  68
FT                   /note= "Wild type Trp is substituted by Glu"
FT   Misc-difference  84
FT                   /note= "Wild type Gln is substituted by Trp"
FT   Misc-difference  95
FT                   /note= "Wild type Ala is substituted by Pro"
FT   Misc-difference  97
FT                   /note= "Wild type Lys is substituted by Cys"
FT   Misc-difference  168
FT                   /note= "Wild type Ser is substituted by Glu"
FT   Misc-difference  181
FT                   /note= "Wild type Arg is substituted by Tyr"
FT   Misc-difference  226
FT                   /note= "Wild type Asn is substituted by Cys"
FT   Misc-difference  277
FT                   /note= "Wild type Tyr is substituted by Asp"
FT   Misc-difference  284
FT                   /note= "Encoded by GG"
FT   Misc-difference  318
FT                   /note= "Encoded by CA"
FT   Misc-difference  351
FT                   /note= "Encoded by CA"
FT   Misc-difference  357
FT                   /note= "Encoded by TT"
XX
PN   US2005246780-A1.
XX
PD   03-NOV-2005.
XX
PF   11-FEB-2005; 2005US-00056354.
XX
PR   13-AUG-1997;  97US-00910798.
PR   01-MAR-1999;  99US-00259214.
PR   13-APR-1999;  99US-00291931.
PR   25-MAY-1999;  99US-00318528.
PR   25-MAY-2000;  2000US-00580515.
PR   24-MAY-2001;  2001US-00866379.
PR   24-MAY-2002;  2002US-00156660.
XX
PA   (SHOR/)  SHORT J M.
PA   (KRET/)  KRETZ K.
PA   (GRAY/)  GRAY K A.
PA   (BART/)  BARTON N R.
PA   (GARR/)  GARRETT J B.
PA   (ODON/)  ODONOGHUE E.
PA   (MATH/)  MATHUR E J.
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, Odonoghue E;
PI   Mathur EJ;
XX
DR   WPI; 2005-733851/75.
DR   N-PSDB; AED50817, AED50824.
```

```
Db   421  IVNEARIPACSL  432
          ||||||||||||
          421  IVNEARIPACSL  432
```

XX
PT New isolated or recombinant nucleic acid encoding phytase enzymes, useful
PT as foodstuff, and for oil degumming, producing an animal feed, delivering
PT a phytase enzyme supplement to an animal.
XX
PS Claim 181; SEQ ID NO 2; 104pp; English.
XX
CC The present invention relates to a recombinant phytase (EC 3.1.2.8)
CC protein and its encoding also nucleic acid molecule. Phytases of the
CC invention is a thermostable protein. They are useful in foodstuffs, for
CC oil degumming, producing an animal feed, delivering a phytase enzyme
CC supplement to an animal, increasing the resistance of a phytase
CC polypeptide to enzymatic inactivation in a digestive system of an animal
CC and in processing of corn and sorghum kernels. The invention is also
CC useful in gene therapy and in production of transgenic animals. The
CC present sequence is a Escherichia coli phytase mutant protein which is
CC encoded by appA gene.
XX
SQ Sequence 432 AA;

Query Match   Best Local Similarity 100.0%; Score 2258; DB 9; Length 432;
Matches 432; Conservative   0; Mismatches   0; Indels   0; Gaps   0;
100.0%;   Pred. No. 3.2e-219;

```
Qy          1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60

Qy         61  TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTKTGTE  120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61  TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTKTGTE  120

Qy        121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180

Qy        181  YQTAFRELERVLNFPQSNLCLKREBKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT  240
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181  YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT  240

Qy        241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA  300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA  300

Qy        301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360

Qy        361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420

Qy        421  IVNEARIPACSL  432
               ||||||||||||
Db        421  IVNEARIPACSL  432
```

RESULT 3
AEE75421
ID   AEE75421 standard; protein; 432 AA.
XX
AC   AEE75421;
XX
DT   23-FEB-2006 (first entry)
XX
DE   E. coli B modified appA phytase 819PH59 SEQ ID NO: 10.
XX
KW   enzyme; mutein; thermostable; phytase; protein stabilization;
KW   pharmaceutical; fertilizer; cramp; cell culture; osteoporosis;
KW   osteopathic; food; appA.
XX
OS   Escherichia coli K12.
OS   Synthetic.

XX
FH   Key   Location/Qualifiers
FT   Misc-difference 68
FT   /note= "Wild type residue Gln replaced by Trp"
FT   Misc-difference 72
FT   /note= "Encoded by CGN"
FT   Misc-difference 84
FT   /note= "Wild type residue Gln replaced by Trp"
FT   Misc-difference 95
FT   /note= "Wild type residue Ala replaced by Pro"
FT   Misc-difference 97
FT   /note= "Wild type residue Lys replaced by Cys"
FT   Misc-difference 168
FT   /note= "Wild type residue Ser replaced by Glu"
FT   Misc-difference 181
FT   /note= "Wild type residue Arg replaced by Tyr"
FT   Misc-difference 226
FT   /note= "Wild type residue Asn replaced by Cys"
FT   Misc-difference 277
FT   /note= "Wild type residue Tyr replaced by Asp"
XX
PN   US2005281792-A1.
XX
PD   22-DEC-2005.
XX
PF   01-SEP-2004; 2004US-0093115.
XX
PR   13-AUG-1997; 97US-00910798.
PR   01-MAR-1999; 99US-00259214.
PR   13-APR-1999; 99US-00291931.
PR   25-MAY-1999; 99US-00318528.
PR   25-MAY-2000; 2000US-00580515.
PR   24-MAY-2001; 2001US-00866379.
XX
PA   (SHOR/) SHORT J M.
PA   (KRET/) KRETZ K A.
PA   (GRAY/) GRAY K A.
PA   (BART/) BARTON N R.
PA   (GARR/) GARRETT J B.
PA   (ODON/) ODONOGHUE E.
PA   (BAUM/) BAUM W.
PA   (ROBE/) ROBERTSON D E.
PA   (ZORN/) ZORNER P.
XX
PI   Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, Odonoghue E;
PI   Baum W, Robertson DE, Zorner P;
XX
DR   WPI; 2006-055022/06.
DR   N-PSDB; AEE75420.
XX
PT   Formulation useful as dietary supplement for treating, preventing or
PT   reversing osteoporosis and bone loss, and preventing muscle cramps,
PT   comprises one or more polypeptides having phytase activity.
XX
PS   Claim 1; SEQ ID NO 10; 82pp; English.
XX
CC   The present sequence is that a novel modified phytase derived from the
CC   Escherichia coli appA protein, with improved thermal tolerance and
CC   protease stability compared to the wild-type. The present invention
CC   relates to a novel formulation, comprising at least one polypeptide
CC   having phytase activity, where the polypeptide is a modified sequence
CC   (AEE75421) derived from the Escherichia coli K-12 appA phytase
CC   (AEE75419). The modification of the enzyme, comprising a series of
CC   mutations, improves the thermal tolerance and protease stability of the
CC   protein. The specification also claims a pharmaceutical composition; a
CC   kit; an immobilized phytase; a fertilizer or soil additive; a liquid
CC   supplement for preventing muscle cramps; a hydrating agent; a tissue
CC   culture or cell culture media; and a plant food additive, all comprising
CC   the novel phytase, and a method of reducing pollution and increasing
CC   nutrient availability in an environment or environmental sample by
CC   degrading environmental phytic acid. The novel enzyme is a dietary
CC   supplement useful for treating, preventing or reversing osteoporosis or
CC   bone loss, and preventing muscle cramps. The liquid supplement is useful

CC for preventing muscle cramps. The method is useful for reducing pollution
CC and increasing nutrient availability in an environment or environmental
CC sample by degrading environmental phytic acid, where the environment or
CC environmental sample comprises a soil or a body of water. The immobilized
CC phytase is useful in foodstuffs for improving the feeding value of
CC phytate rich ingredients.
CC
XX
SQ Sequence 432 AA;

Qy Query Match
Best Local Similarity   100.0%; Score 2258; DB 10; Length 432;
Matches 432; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP   60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP   60

Qy 61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120

Qy 121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180

Qy 181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240

Qy 241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300

Qy 301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGGELVFERW 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGGELVFERW 360

Qy 361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLANTPPGEVKLTLLAGCEERNAQGMCSLAGFTQ 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLANTPPGEVKLTLLAGCEERNAQGMCSLAGFTQ 420

Qy 421 IVNEARIPACSL 432
        ||||||||||||
Db 421 IVNEARIPACSL 432

RESULT 4
ID   ADC87743 standard; protein; 436 AA.
XX
AC   ADC87743;
XX
DT   01-JAN-2004 (first entry)
XX
DE   Modified Escherichia coli phytase.
XX
KW   Phytase; food supplement; enzyme delivery matrix; soybean meal;
KW   thermotolerance; thermostability; kernel; phytate;
KW   myo-inositol-hexaphosphate; inositol; inorganic phosphate;
KW   thermotolerant; feed value; digestion; enzyme.
XX
OS   Synthetic.
OS   Escherichia coli.
XX
PN   US2003103958-A1.
XX
PD   05-JUN-2003.
XX
PF   24-MAY-2002; 2002US-00156660.
XX
PR   13-AUG-1997; 97US-00910798.
PR   01-MAR-1999; 99US-0025214.
PR   13-APR-1999; 99US-00291931.

PA   (DIVE-) DIVERSA CORP.
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI   Mathur EJ;
XX
DR   WPI; 2003-787039/74.
DR   N-PSDB; ADC87742.
XX
PT   New nucleic acid encoding a polypeptide having phytase activity, useful
PT   in improving the feeding value of phytate rich ingredients or as an aid
PT   in phytate digestion.
XX
PS   Claim 60; SEQ ID NO 2; 113pp; English.
XX
CC   The invention discloses a new isolated or recombinant nucleic acid which
CC   encodes a polypeptide having a phytase activity. Also claimed is a
CC   nucleic acid probe, an amplification primer sequence pair, an expression
CC   cassette comprising the nucleic acid, a vector comprising the nucleic
CC   acid, a transgenic non-human animal or plant, or its seed, comprising the
CC   nucleic acid, an antisense oligonucleotide, inhibiting the translation of
CC   a phytase message in a cell, a heterodimer comprising the polypeptide and
CC   a second domain, an array comprising immobilised polypeptide or nucleic
CC   acid, a hybridoma comprising an antibody that specifically binds to the
CC   polypeptide, a food supplement for an animal, an edible enzyme delivery
CC   matrix, an edible pellet comprising a granule edible carrier and the
CC   polypeptide, a feed composition, a soybean meal, isolating or identifying
CC   the polypeptide, making an anti-phytase antibody, producing a recombinant
CC   polypeptide, determining whether a compound binds to the polypeptide,
CC   identifying a modulator, whole cell engineering of new or modified
CC   phenotypes by using real-time metabolic flux analysis, increasing
CC   thermotolerance or thermostability of the phytase polypeptide, increasing
CC   the resistance of the phytase polypeptide to enzymatic inactivation in a
CC   digestive system and processing of corn and sorghum kernels. The phytase
CC   activity comprises catalysis of phytate (myo-inositol-hexaphosphate) to
CC   inositol and inorganic phosphate and is thermotolerant. The nucleic acid
CC   is useful in improving the feeding value of phytate rich ingredients or
CC   as an aid in phytate digestion. The sequence presented is the modified
CC   Escherichia coli phytase.
XX
SQ   Sequence 436 AA;

Qy Query Match
Best Local Similarity   100.0%; Score 2258; DB 7; Length 436;
Matches 432; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP   60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP   60

Qy 61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120

Qy 121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180

Qy 181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240

Qy 241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300

Qy 301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGGELVFERW 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGGELVFERW 360

```
Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy  421 IVNEARIPACSL 432
        ||||||||||||
Db  421 IVNEARIPACSL 432
```

RESULT 5
AAE15808
ID   AAE15808 standard; protein; 430 AA.
XX
AC   AAE15808;
XX
DT   26-MAR-2002 (first entry)
XX
DE   Escherichia coli appA phytase mutant protein.
XX
KW   Bacterial phytase; K12 appA phytase; protease stability; anabolic;
KW   gastrointestinal; nutritional value; feed treatment process; therapy;
KW   thermal tolerance; growth performance; alcoholic drink; biopulping;
KW   non-alcoholic drink; biobleaching; mutant; mutein.
XX
OS   Escherichia coli.
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FH
FT   Misc-difference 68
FT   /note= "Wild type Trp substituted with Glu"
FT   Misc-difference 84
FT   /note= "Wild type Gln substituted with Trp"
FT   Misc-difference 95
FT   /note= "Wild type Ala substituted with Pro"
FT   Misc-difference 97
FT   /note= "Wild type Lys substituted with Cys"
FT   Misc-difference 168
FT   /note= "Wild type Ser substituted with Glu"
FT   Misc-difference 180
FT   /note= "Wild type Arg substituted with Tyr"
FT   Misc-difference 225
FT   /note= "Wild type Asn substituted with Cys"
FT   Misc-difference 276
FT   /note= "Wild type Tyr substituted with Asp"

WO200190333-A2.
XX
PN   29-NOV-2001.
XX
PD   24-MAY-2001; 2001WO-US017118.
XX
PR   25-MAY-2000; 2000US-00580515.
XX
PA   (DIVE-) DIVERSA CORP.
XX
PI   Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, O'donoghue E;
XX
DR   WPI; 2002-083108/11.
XX
PT   Claim 59; Fig 8; 170pp; English.

New bacterial phytase for e.g. improving the nutritional value of phytate
-containing foodstuffs and subsequently improving the growth performance
of an organism that consumes it, or in treating animal digestive systems.

The patent discloses recombinant bacterial phytase from Escherichia coli
K12 appA phytase. The enzyme has phytase activity and improved thermal
tolerance when compared with wild-type phytase. It has improved protease
stability at low pH. The recombinant phytase is useful for improving the
nutritional value of phytate-containing foodstuffs and subsequently
improving the growth performance of an organism that consumes it, in
treating animal digestive systems, in feed treatment processes and for in
vitro purposes related to research, discovery and development. They are

CC   also used for generating recombinant digestive system life forms, for
CC   producing or manufacturing alcoholic and non-alcoholic drinks based on
CC   the use of moulds, grains and/or plants, in biopulping and bio-
CC   bleaching where a reduction in the use of environmentally harmful chemicals that
CC   are traditionally used in the pulp and paper industry is desired and in
CC   the reduction or possible elimination of the need for mineral
CC   supplements, enzymes or therapeutic drugs for animals from the daily feed
CC   thus increasing the amount calories present and nutrients present in the feed.
CC   The present sequence is E. coli appA phytase mutant protein
XX
SQ   Sequence 430 AA;

Query Match   98.7%; Score 2228; DB 5; Length 430;
Best Local Similarity   99.5%; Pred. No. 3.5e-216;
Matches 430; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

```
Qy    1 MKAILIPFLSLLIPLTTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTKKTGE 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTKKTGE 120

Qy  121 APAAGLAPDCAITVHTQADTSSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
        |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db  121 APAAGLAPDCAITVHTQADTSSSPDPLFNPLKTGVCQLDNANVTDAILE-AGGSIADFTGH 179

Qy  181 YQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  180 YQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 239

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
        |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db  240 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFD-LQRTPEVARSRATPLLDLIKTA 298

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANIGGALELANWTLPGQPDNTPGGELVFERW 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  299 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANIGGALELANWTLPGQPDNTPGGELVFERW 358

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  359 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 418

Qy  421 IVNEARIPACSL 432
        ||||||||||||
Db  419 IVNEARIPACSL 430
```

RESULT 6
ADA19451
ID   ADA19451 standard; protein; 430 AA.
XX
AC   ADA19451;
XX
DT   20-NOV-2003 (first entry)
XX
DE   E. coli K12 phytase mutant.
XX
KW   Phytase; enzyme; phytate; appA gene; animal feed; inorganic phosphate;
KW   digestion enhancement; transgenic; thermal tolerance; protease stability;
KW   mutant; mutein.
XX
OS   Escherichia coli; strain K12.
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FH
FT   Misc-difference 68
FT   /note= "Wild-type Trp substituted by Glu"
FT   Misc-difference 72
FT   /note= "Encoded by CGN"
FT   Misc-difference 84
FT   /note= "Wild-type Gln substituted by Trp"

FT   Misc-difference 95
FT   /note= "Wild-type Ala substituted by Pro"
FT   Misc-difference 97
FT   /note= "Wild-type Lys substituted by Cys"
FT   Misc-difference 168
FT   /note= "Wild-type Ser substituted by Glu"
FT   Misc-difference 180
FT   /note= "Wild-type Arg substituted by Tyr"
FT   Misc-difference 226
FT   /note= "Wild-type Asn substituted by Cys"
FT   Misc-difference 277
FT   /note= "Wild-typr Tyr substituted by Asp"
XX
PN   US20020136754-A1.
XX
PD   26-SEP-2002.
XX
PF   24-MAY-2001; 2001US-00866379.
XX
PR   13-AUG-1997; 97US-00910798.
PR   01-MAR-1999; 99US-00259214.
PR   13-APR-1999; 99US-00291931.
PR   25-MAY-1999; 99US-00318528.
PR   25-MAY-2000; 2000US-00580515.
XX
PA   (SHOR/) SHORT J M.
PA   (KRET/) KRETZ K A.
PA   (GRAY/) GRAY K A.
PA   (BART/) BARTON N R.
PA   (GARR/) GARRETT J B.
PA   (DONO/) O' DONOGHUE E.
PA   (MATH/) MATHUR E J.
XX
PI   Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB;
PI   O' Donoghue E, Mathur EJ;
XX
DR   WPI; 2003-040002/03.
XX

PS   Claim 59; Fig 8; 62pp; English.
XX

CC   Isolated Escherichia coli polynucleotide encoding a modified phytase
CC   enzyme, useful in the production of animal feed, for improving the
CC   nutritional value of phytate-containing foodstuff and for enhancing
CC   digestion in humans and animals.
XX

CC   The invention relates to an isolated Escherichia coli polynucleotide
CC   encoding a phytase enzyme appearing as ADA19450 and having amino acids
CC   modification from W68P, Q84W, A95P, K97C, S168E, R180Y, N226C or Y277D.
CC   Also included the E. coli appA gene ADA19449 (or an oligonucleotide
CC   derived from it) or its mutant sequence ADA19452, expression vectors,
CC   host cells, a method of improving nutritional value of a phytate-
CC   containing foodstuff by contacting the phytate-containing foodstuff with
CC   a pure phytase enzyme mutated phytase (where the phytase enzyme catalyzes
CC   the liberation of inorganic phosphate from the phytate in the phytate-
CC   containing foodstuff), a method to produce an animal feed containing a
CC   microbial phytase (comprising culturing the plant cell, plant part or
CC   plant under conditions where the nucleotide sequence is expressed and
CC   converting the plant cells, plants or plant into a composition for animal
CC   feed), a feed composition for animals (comprising the plant seeds, plant
CC   cells, plant parts or plants in admixture with a phytate-containing
CC   foodstuff), a method to treat a human or an animal able to benefit from
CC   digestive enhancement by the activity of an exogenous phytase enzymes
CC   comprising administering to the human or animal the plant seed, plant
CC   cells, plant parts or plants of a transgenic plant which is modified to
CC   contain an expression system which expresses a nucleotide sequence
CC   encoding a phytase enzyme, a transgenic non-human organism whose genome
CC   comprising a heterologous nucleic acid sequence encoding a polypeptide
CC   having phytase activity. The phytase enzyme is useful for improving the
CC   nutritional value of phytate-containing foodstuff, in the production of
CC   animal feed and for enhancing digestion in humans and animals. The
CC   invented method improves thermal tolerance and protease stability. It
CC   also improves the feeding value of phytate rich ingredients. The present
CC   sequence represents E. coli K12 mutant phytase.

XX
SQ   Sequence 430 AA;
XX

Query Match   98.7%; Score 2228; DB 6; Length 430;
Best Local Similarity   99.5%; Pred. No. 3.5e-216;
Matches 430; Conservative   0; Mismatches   0; Indels   2; Gaps   2;

Qy   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120

Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||| -|||||||
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILE-AGGSIADFTGH 179

Qy 181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 180 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 239

Qy 241 EIFLLQQAQGMPEPGWGRRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
         |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db 240 EIFLLQQAQGMPEPGWGRRITDSHQWNTLLSLHNAQFD-LQRTPEVARSRATPLLDLIKTA 298

Qy 301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 299 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 358

Qy 361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 359 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 418

Qy 421 IVNEARIPACSL 432
         |||||||||||
Db 419 IVNEARIPACSL 430

RESULT 7
AAB36257
ID   AAB36257 standard; protein; 432 AA.
XX
AC   AAB36257;
XX
DT   12-SEP-2003 (revised)
DT   20-FEB-2001 (first entry)
XX
DE   Lama2/APPA plasmid translated sequence.
XX
KW   Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW   environmental pollution; pig.
XX
OS   Mus musculus.
OS   Escherichia coli.
OS   Chimeric.
XX
PN   WO200064247-A1.
XX
PD   02-NOV-2000.
XX
PF   20-APR-2000; 2000WO-CA000430.
XX
PR   23-APR-1999; 99US-0130508P.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   Forsberg CW, Golovan S, Phillips JP;
XX
DR   WPI; 2000-687245/67.
DR   N-PSDB; AAC68294.
XX

Left column

```
PT   Transgenic non-human animal for gastrointestinal tract specific
PT   expression of a protein, preferably phytase, comprises a nucleic acid
PT   sequence including a heterologous transgene construct encoding the
PT   protein.
XX
PS   Disclosure; Pig 5; 152pp; English.
XX
CC   The present invention provides transgenic animals which produce desired
CC   proteins, in this case pigs which expresses phytase in the salivary
CC   gland. Low phytase production levels result in phytate in the diet being
CC   excreted and causing phosphorus contamination in water, as well as
CC   reducing the growth of animals. The invention provides a number of
CC   transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC   on 12-SEP-2003 to standardise OS field)
XX
SQ   Sequence 432 AA;

Query Match             96.6%;  Score 2182;  DB 3;  Length 432;
Best Local Similarity   98.1%;  Pred. No. 1.6e-211;
Matches   424;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      1  MKAILIPFLSLIIPLITPQSAPAQGSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60
           |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db      1  MKAILIPFLSLIIPLITPQSAFAQGSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60

Qy     61  TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE  120
           |||||| |||||| ||||||||| |||||||||  |||||||||||||||||||||||||
Db     61  TWPVKLGWLTPRGGBLIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120

Qy    121  APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFGH   180
           |||||||||||||||||||||||||||||||||||||||||||||||| |||||||| ||
Db    121  APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180

Qy    181  RQTAFRELERVLNFPQSNLCLRBEKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
           ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    181  RQTAFRELERVLNFPQSNLCLRBEKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT  240

Qy    241  EIFLLQOAQGMPEPGWGRRITDSHQWNTLLSLHNAQPDLLQRTPEVARSRATPLLDLIKTA 300
           |||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db    241  EIFLLQOAQGMPEPGWGRRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy    301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360

Qy    361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy    421  IVNEARIPACSL 432
           ||||||||||||
Db    421  IVNEARIPACSL 432

RESULT 8
AAB36261
ID   AAB36261 standard; protein; 432 AA.
XX
AC   AAB36261;
XX
DT   12-SEP-2003 (revised)
DT   20-FEB-2001 (first entry)
XX
DE   R15/APPA plasmid translated sequence.
XX
KW   Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW   environmental pollution; pig.
XX
OS   Rattus sp.
OS   Escherichia coli.
OS   Chimeric.
XX
PN   WO20006424247-A1.
```

Right column

```
XX
PD   02-NOV-2000.
XX
PD   20-APR-2000; 2000WO-CA000430.
XX
PR   23-APR-1999; 99US-0130508P.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   Forsberg CW, Golovan S, Phillips JP;
XX
DR   WPI; 2000-687245/67.
DR   N-PSDB; AAC68298.
XX
PT   Transgenic non-human animal for gastrointestinal tract specific
PT   expression of a protein, preferably phytase, comprises a nucleic acid
PT   sequence including a heterologous transgene construct encoding the
PT   protein.
XX
PS   Disclosure; Fig 21; 152pp; English.
XX
CC   The present invention provides transgenic animals which produce desired
CC   protein, in this case pigs which expresses phytase in the salivary
CC   gland. Low phytase production levels result in phytate in water,
CC   excreted and causing phosphorus contamination in water, as well as
CC   reducing the growth of animals. The invention provides a number of
CC   transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC   on 12-SEP-2003 to standardise OS field)
XX
SQ   Sequence 432 AA;

Query Match             96.6%;  Score 2182;  DB 3;  Length 432;
Best Local Similarity   98.1%;  Pred. No. 1.6e-211;
Matches   424;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      1  MKAILIPFLSLIIPLITPQSAPAQGSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60
           |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db      1  MKAILIPFLSLIIPLITPQSAFAQGSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60

Qy     61  TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE  120
           |||||| |||||| ||||||||| |||||||||  |||||||||||||||||||||||||
Db     61  TWPVKLGWLTPRGGBLIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120

Qy    121  APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFGH   180
           |||||||||||||||||||||||||||||||||||||||||||||||| |||||||| ||
Db    121  APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180

Qy    181  RQTAFRELERVLNFPQSNLCLRBEKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
           ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    181  RQTAFRELERVLNFPQSNLCLRBEKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT  240

Qy    241  EIFLLQOAQGMPEPGWGRRITDSHQWNTLLSLHNAQPDLLQRTPEVARSRATPLLDLIKTA 300
           |||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db    241  EIFLLQOAQGMPEPGWGRRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy    301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360

Qy    361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy    421  IVNEARIPACSL 432
           ||||||||||||
Db    421  IVNEARIPACSL 432

RESULT 9
AAB36262
ID   AAB36262 standard; protein; 432 AA.
XX
AC   AAB36262;
```

```
XX
DT   12-SEP-2003  (revised)
DT   20-FEB-2001  (first entry)
XX
DE   SV40/APPA plasmid translated sequence.
XX
KW   Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW   environmental pollution; pig.
XX
OS   Simian virus 40.
OS   Escherichia coli.
OS   Chimeric.
XX
PN   WO200064247-A1.
XX
PD   02-NOV-2000.
XX
PF   20-APR-2000; 2000WO-CA000430.
XX
PR   23-APR-1999;  99US-0130508P.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   Forsberg CW,  Golovan S,  Phillips JP;
XX
DR   WPI; 2000-687245/67.
DR   N-PSDB; AAC68299.
XX
PT   Transgenic non-human animal for gastrointestinal tract specific
PT   expression of a protein, preferably phytase, comprises a nucleic acid
PT   sequence including a heterologous transgene construct encoding the
PT   protein.
XX
PS   Disclosure; Fig 22; 152pp; English.
XX
CC   The present invention provides transgenic animals which produce desired
CC   proteins, in this case pigs which expresses phytase in the salivary
CC   gland. Low phytase production levels result in phytate in the diet being
CC   excreted and causing phosphorus contamination in water, as well as
CC   reducing the growth of animals. The invention provides a number of
CC   transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC   on 12-SEP-2003 to standardise OS field)
XX
SQ   Sequence 432 AA;

Query Match             96.6%;  Score 2182;  DB 3;  Length 432;
Best Local Similarity   98.1%;
Matches  424;  Conservative   0;  Mismatches   8;  Indels    0;  Gaps    0;

Qy          1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy         61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TWPVKLGELTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy        121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy        181 ROTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240

Qy        241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLLDLIKTA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLLDLIKTA 300

Qy        301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQDNTPPGGELVPERW 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LTPHPPQKQAVGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360

Qy        361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLANTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLANTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy        421 IVNEARIPACSL 432
              ||||||||||||
Db        421 IVNEARIPACSL 432
```

```
RESULT 10
ID   AAB36259 standard; protein; 432 AA.
XX
AC   AAB36259;
XX
DT   12-SEP-2003  (revised)
DT   20-FEB-2001  (first entry)
XX
DE   R15/APPA plasmid translated sequence.
XX
KW   Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW   environmental pollution; pig.
XX
OS   Rattus sp.
OS   Escherichia coli.
OS   Chimeric.
XX
PN   WO200064247-A1.
XX
PD   02-NOV-2000.
XX
PF   20-APR-2000; 2000WO-CA000430.
XX
PR   23-APR-1999;  99US-0130508P.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   Forsberg CW,  Golovan S,  Phillips JP;
XX
DR   WPI; 2000-687245/67.
DR   N-PSDB; AAC68296.
XX
PT   Transgenic non-human animal for gastrointestinal tract specific
PT   expression of a protein, preferably phytase, comprises a nucleic acid
PT   sequence including a heterologous transgene construct encoding the
PT   protein.
XX
PS   Disclosure; Fig 19; 152pp; English.
XX
CC   The present invention provides transgenic animals which produce desired
CC   proteins, in this case pigs which expresses phytase in the salivary
CC   gland. Low phytase production levels result in phytate in the diet being
CC   excreted and causing phosphorus contamination in water, as well as
CC   reducing the growth of animals. The invention provides a number of
CC   transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC   on 12-SEP-2003 to standardise OS field)
XX
SQ   Sequence 432 AA;

Query Match             96.6%;  Score 2182;  DB 3;  Length 432;
Best Local Similarity   98.1%;
Matches  424;  Conservative   0;  Mismatches   8;  Indels    0;  Gaps    0;

Qy          1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy         61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TWPVKLGELTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy        121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
```

```
Qy        361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLANTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLANTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy        421 IVNEARIPACSL 432
              ||||||||||||
Db        421 IVNEARIPACSL 432
```

RESULT 11

```
AAB36258
ID   AAB36258 standard; protein; 432 AA.
XX
AC   AAB36258;
XX
DT   12-SEP-2003  (revised)
DT   20-FEB-2001  (first entry)
XX
DE   R15/APPA plasmid translated sequence.
XX
KW   Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW   environmental pollution; pig.
XX
OS   Rattus sp.
OS   Escherichia coli.
OS   Chimeric.
XX
PN   WO20006424247-A1.
XX
PD   02-NOV-2000.
XX
PF   20-APR-2000; 2000WO-CA000430.
XX
PR   23-APR-1999;  99US-0130508P.
XX
PA   (UYGU-) UNIV GUELPH.
PI   Forsberg CW, Golovan S, Phillips JP;
XX
DR   WPI; 2000-687245/67.
DR   N-PSDB; AAC68295.
XX
PS   Disclosure; Pig 18; 152pp; English.
XX
PT   Transgenic non-human animal for gastrointestinal tract specific
PT   expression of a protein, preferably phytase, comprises a nucleic acid
PT   sequence including a heterologous transgene construct encoding the
PT   protein.
XX
CC   The present invention provides transgenic animals which produce desired
CC   proteins, in this case pigs which expresses phytase in the salivary
CC   gland. Low phytase production levels result in phytate in the diet being
CC   excreted and causing phosphorus contamination in water, as well as
CC   reducing the growth of animals. The invention provides a number of
CC   transgenes containing the B. coli APPA phytase coding sequence. (Updated
CC   on 12-SEP-2003 to standardise OS field)
XX
SQ   Sequence 432 AA;
```

```
Db   121  APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180
Qy   181  YQTAFRELERVLNFPQOSNLCLRKEKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT  240
          | |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   181  RQTAFRELERVLNFPQOSNLCLRKEKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
Qy   241  EIFPLLQQAQGMPEPGWGNGRITDSHQWNTLLSLRHNAQPDLLQRTPEVARSRATPLLDLIKTA  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  EIFPLLQQAQGMPEPGWGNGRITDSHQWNTLLSLRHNAQPDLLQRTPEVARSRATPLLDLIKTA  300
Qy   301  LTPHPPQOKQAYGYTLPTSVLFIAGHDTNLANLGGALELNWTLPCQPDNTPPGGELVFERW  360
          ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   301  LTPHPPQOKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPCQPDNTPPGGELVFERW  360
Qy   361  RRLSDNSQWIQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  RRLSDNSQWIQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
Qy   421  IVNEARIPACSL  432
          ||||||||||||
Db   421  IVNEARIPACSL  432
```

RESULT 12

```
AAB36263
ID   AAB36263 standard; protein; 432 AA.
XX
AC   AAB36263;
XX
DT   12-SEP-2003  (revised)
DT   20-FEB-2001  (first entry)
XX
DE   Lama2/APPA plasmid translated sequence.
XX
KW   Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW   environmental pollution; pig.
XX
OS   Mus musculus.
OS   Escherichia coli.
OS   Chimeric.
XX
PN   WO20006424247-A1.
XX
PD   02-NOV-2000.
XX
PF   20-APR-2000; 2000WO-CA000430.
XX
PR   23-APR-1999;  99US-0130508P.
XX
PA   (UYGU-) UNIV GUELPH.
PI   Forsberg CW, Golovan S, Phillips JP;
XX
DR   WPI; 2000-687245/67.
DR   N-PSDB; AAC68300.
XX
PT   Transgenic non-human animal for gastrointestinal tract specific
PT   expression of a protein, preferably phytase, comprises a nucleic acid
PT   sequence including a heterologous transgene construct encoding the
```

```
Query Match    96.6%;  Score 2182;  DB 3;  Length 432;
Best Local Similarity  98.1%;  Pred. No.1.6e-211;
Matches  424;  Conservative  0;  Mismatches  8;  Indels  0;  Gaps  0;

Qy   1    MKAILIPFLSLLILPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1    MKAILIPFLSLLILPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP  60
Qy   61   TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQOSGQVAIIADVDERTRKTGE  120
          |||| || |||||||||||||| |||||||||||| ||||||||||||||||||||||||
Db   61   TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQOSGQVAIIADVDERTRKTGE  120
Qy   121  APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180
Qy   181  RQTAFRELERVLNFPQOSNLCLRKEKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  RQTAFRELERVLNFPQOSNLCLRKEKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
Qy   241  EIFPLLQQAQGMPEPGWGNGRITDSHQWNTLLSLRHNAQFDLLQRTPEVARSRATPLLDLIKTA  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  EIFPLLQQAQGMPEPGWGNGRITDSHQWNTLLSLRHNAQFDLLQRTPEVARSRATPLLDLIKTA  300
Qy   301  LTPHPPQOKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPCQPDNTPPGGELVFERW  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  LTPHPPQOKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPCQPDNTPPGGELVFERW  360
Qy   361  RRLSDNSQWIQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  RRLSDNSQWIQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
Qy   421  IVNEARIPACSL  432
          ||||||||||||
Db   421  IVNEARIPACSL  432
```

PT   protein.
XX
PS   Disclosure; Fig 23; 152pp; English.
XX
CC   The present invention provides transgenic animals which produce desired
CC   proteins, in this case pigs which expresses phytase in the salivary
CC   gland. Low phytase production levels result in phytate in the diet being
CC   excreted and causing phosphorus contamination in water, as well as
CC   reducing the growth of animals. The invention provides a number of
CC   transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC   on 12-SEP-2003 to standardise OS field)
XX
SQ   Sequence 432 AA;

Query Match   96.6%; Score 2182; DB 3; Length 432;
Best Local Similarity   98.1%;
Matches 424; Conservative   0; Mismatches   8; Indels   0; Gaps   0;

```
Qy          1 MKAILIPFLSLLILPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKAILIPFLSLLILPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy         61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy        121 APFAAGLLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APFAAGLLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy        181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy        241 EIFLLQQAQGMPEPGWNGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 EIFLLQQAQGMPEPGWNGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy        301 LTPHPPQOKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LTPHPPQOKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW 360

Qy        361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy        421 IVNEARIPACSL 432
              ||||||||||||
Db        421 IVNEARIPACSL 432
```

RESULT 13
AC   AAB36260
ID   AAB36260 standard; protein; 432 AA.
XX
AC   AAB36260;
XX
DT   12-SEP-2003 (revised)
DT   20-FEB-2001 (first entry)
XX
DE   R15/APPA plasmid translated sequence.
XX
KW   Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW   environmental pollution; pig.
XX
OS   Rattus sp.
OS   Escherichia coli.
OS   Chimeric.
XX
PN   WO200064247-A1.
XX
PD   02-NOV-2000.
XX

PF   20-APR-2000; 2000WO-CA000430.
XX
PR   23-APR-1999; 99US-0130508P.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   Forsberg CW, Golovan S, Phillips JP;
XX
DR   WPI; 2000-687245/67.
DR   N-PSDB; AAC68297.
XX
PT   Transgenic non-human animal for gastrointestinal tract specific
PT   expression of a protein, preferably phytase, comprises a nucleic acid
PT   sequence including a protein, preferably phytase, comprises a heterologous transgene construct encoding the
PT   protein.
XX
PS   Disclosure; Fig 20; 152pp; English.
XX
CC   The present invention provides transgenic animals which produce desired
CC   proteins, in this case pigs which expresses phytase in the salivary
CC   gland. Low phytase production levels result in phytate in the diet being
CC   excreted and causing phosphorus contamination in water, as well as
CC   reducing the growth of animals. The invention provides a number of
CC   transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC   on 12-SEP-2003 to standardise OS field)
XX
SQ   Sequence 432 AA;

Query Match   96.6%; Score 2182; DB 3; Length 432;
Best Local Similarity   98.1%;
Matches 424; Conservative   0; Mismatches   8; Indels   0; Gaps   0;

```
Qy          1 MKAILIPFLSLLILPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKAILIPFLSLLILPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy         61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy        121 APFAAGLLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 APFAAGLLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy        181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy        241 EIFLLQQAQGMPEPGWNGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 EIFLLQQAQGMPEPGWNGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy        301 LTPHPPQOKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LTPHPPQOKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW 360

Qy        361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy        421 IVNEARIPACSL 432
              ||||||||||||
Db        421 IVNEARIPACSL 432
```

RESULT 14
AC   AAU77775
ID   AAU77775 standard; protein; 432 AA.
XX
AC   AAU77775;
XX
DT   05-JUN-2002 (first entry)
XX

DE   Phytase associated protein.
XX
KW   Phytase.
XX
OS   Unidentified.
XX
PN   KR9906028-A.
XX
PD   15-DEC-1999.
XX
PF   25-MAY-1998; 98KR-0001810.
XX
PR   25-MAY-1998; 98KR-0001810.
XX
PA   (WOOJ-) WOOJIN CO LTD.
XX
PI   Bae HD, Forceburgh CW, Goloben S, Cheng KJ;
XX
DR   WPI; 2000-645078/62.
DR   N-PSDB; ABK12514.
XX
PT   Novel phytase gene, recombinant phytase and usage thereof.
XX
PS   Disclosure; Fig 3; 10pp; Korean.
XX
CC   The invention relates to a novel phytase gene, a recombinant phytase gene
CC   and their uses. This is the amino acid sequence of the phytase associated
CC   protein described in the invention
XX
SQ   Sequence 432 AA;

Query Match   96.6%; Score 2182; DB 3; Length 432;
Best Local Similarity   98.1%;
Matches 424; Conservative   0; Mismatches   8; Indels   0; Gaps   0;

```
Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
        |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy  181 YQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
        |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db  181 RQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLLIKTA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLLIKTA 300

Qy  301 LTPHPQOKQAYGVTLPTSVLFTAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPQOKQAYGVTLPTSVLFTAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLANTPPGEVKLTLAGCEERNAOQGMCSLAGFTQ 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLANTPPGEVKLTLAGCEERNAOQGMCSLAGFTQ 420

Qy  421 IVNEARIPACSL 432
        ||||||||||||
Db  421 IVNEARIPACSL 432
```

XX
DT   26-MAR-2002 (first entry)
XX
DE   Escherichia coli appA phytase wild type protein.
XX
KW   Bacterial phytase; K12 appA phytase; protease stability; anabolic;
KW   gastrointestinal; nutritional value; feed treatment process; therapy;
KW   thermal tolerance; growth performance; alcoholic drink; biopulping;
KW   non-alcoholic drink; biobleaching.
XX
OS   Escherichia coli.
XX
PN   WO200190333-A2.
XX
PD   29-NOV-2001.
XX
PF   24-MAY-2001; 2001WO-US017118.
XX
PR   25-MAY-2000; 2000US-00580515.
XX
PA   (DIVE-) DIVERSA CORP.
XX
PI   Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, O'donoghue E;
XX
DR   WPI; 2002-083108/11.
DR   N-PSDB; AAD25463.
XX
PT   New bacterial phytase for e.g. improving the nutritional value of phytate
PT   -containing foodstuffs and subsequently improving the growth performance
PT   of an organism that consumes it, or in treating animal digestive systems.
XX
PS   Claim 54; Fig 8; 170pp; English.
XX
CC   The patent discloses recombinant bacterial phytase from Escherichia coli
CC   K12 appA phytase. The enzyme has phytase activity and improved thermal
CC   tolerance when compared with wild-type phytase. It has improved protease
CC   stability at low pH. The recombinant phytase is useful for improving the
CC   nutritional value of phytate-containing foodstuffs and subsequently
CC   improving the growth performance of an organism that consumes it, in
CC   treating animal digestive systems, in feed treatment processes and for in
CC   vitro purposes related to research, discovery and development. They are
CC   also used for generating recombinant digestive system life forms, for
CC   producing or manufacturing alcoholic and non-alcoholic drinks based on
CC   the use of moulds, grains and/or plants, in biopulping and bio-bleaching
CC   where a reduction in the use of environmentally harmful chemicals that
CC   are traditionally used in the pulp and paper industry is desired and in
CC   the reduction or possible elimination of the need for mineral
CC   supplements, enzymes or therapeutic drugs for animals from the daily feed
CC   thus increasing the amount calories and nutrients present in the feed.
CC   The present sequence is E. coli appA phytase wild type protein
XX
SQ   Sequence 432 AA;

Query Match   96.6%; Score 2182; DB 5; Length 432;
Best Local Similarity   98.1%;
Matches 424; Conservative   0; Mismatches   8; Indels   0; Gaps   0;

```
Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
        |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy  181 YQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
        |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db  181 RQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
```

RESULT 15
ID   AAE15807 standard; protein 432 AA.
XX
AC   AAE15807;

```
Qy    241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA  300

Qy    301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW  360

Qy    361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLLAGCEERNAQGMCSLAGFTQ  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLLAGCEERNAQGMCSLAGFTQ  420

Qy    421  IVNEARIPACSL  432
           ||||||||||||
Db    421  IVNEARIPACSL  432
```

Search completed: June 13, 2006, 10:22:05
Job time : 138.734 secs

OM protein - protein search, using sw model

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

Run on: June 13, 2006, 10:28:49 ; Search time 35.6697 Seconds
                 (without alignments)
                 1060.093 Million cell updates/sec

Title: US-10-601-319-10
Perfect score: 2258
Sequence: 1 MKAILIPFLSLLIPLTPQSA.........CSLAGFTQIVNEARIPACSL 432
Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5
Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Maximum DB seq length: 200000000
Minimum DB seq length: 0

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued Patents AA:*
                 1: /EMC_Celerra_SIDS3/ptcodata/2/iaa/5_COMB.pep:*
                 2: /EMC_Celerra_SIDS3/ptcodata/2/iaa/6_COMB.pep:*
                 3: /EMC_Celerra_SIDS3/ptcodata/2/iaa/7_COMB.pep:*
                 4: /EMC_Celerra_SIDS3/ptcodata/2/iaa/H_COMB.pep:*
                 5: /EMC_Celerra_SIDS3/ptcodata/2/iaa/PCTUS_COMB.pep:*
                 6: /EMC_Celerra_SIDS3/ptcodata/2/iaa/RB_COMB.pep:*
                 7: /EMC_Celerra_SIDS3/ptcodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2228	98.7	430	2	US-09-866-379-10	Sequence 10, Appli
2	2182	96.6	432	2	US-11-083-452-8	Sequence 8, Appli
3	2177	96.4	432	2	US-09-635-504-33	Sequence 33, Appl
4	2171	96.1	440	2	US-10-062-848-78	Sequence 78, Appl
5	2171	96.1	440	3	US-11-229-358-6	Sequence 6, Appli
6	2171	96.1	440	2	US-09-684-855-107	Sequence 107, App
7	2171	96.1	440	2	US-09-684-855-129	Sequence 129, App
8	2171	96.1	440	2	US-09-580-515-2	Sequence 2, Appli
9	2157	95.5	432	2	US-09-715-477-3	Sequence 3, Appli
10	2147	95.1	432	2	US-09-866-379-2	Sequence 2, Appli
11	2147	95.1	433	2	US-09-540-149A-1	Sequence 1, Appli
12	2063.5	91.4	423	1	US-10-266-041A-1	Sequence 1, Appli
13	543	24.0	421	2	US-09-489-039A-7512	Sequence 7512, Ap
14	381	16.9	522	2	US-09-489-039A-13501	Sequence 13501, A
15	126	5.6	439	2	US-09-044-718-3	Sequence 3, Appli
16	126	5.6	439	2	US-10-062-848-3	Sequence 3, Appli
17	126	5.6	440	2	US-09-684-855-106	Sequence 106, App
18	126	5.6	440	2	US-09-684-855-128	Sequence 128, App
19	126	5.6	440	2	US-09-684-855-151	Sequence 151, App
20	126	5.6	440	2	US-09-488-265B-6	Sequence 6, Appli
21	126	5.6	449	2	US-09-044-718-12	Sequence 12, Appl
22	126	5.6	449	2	US-10-062-848-12	Sequence 12, Appl
23	126	5.6	465	2	US-09-868-435-33	Sequence 33, Appl
24	126	5.6	465	2	US-09-744-231-33	Sequence 33, Appl
25	126	5.6	465	2	US-09-044-718-78	Sequence 78, Appl
26	126	5.6	465	2	US-09-636-499-6	Sequence 6, Appli
27	126	5.6	465	2	US-09-273-871A-8	Sequence 8, Appli
28	126	5.6	465	2	US-11-083-452-8	Sequence 8, Appli
29	126	5.6	465	2	US-09-635-504-33	Sequence 33, Appl
30	126	5.6	465	2	US-10-062-848-78	Sequence 78, Appl
31	126	5.6	465	3	US-11-229-358-6	Sequence 6, Appli
32	126	5.5	440	2	US-09-684-855-107	Sequence 107, App
33	126	5.5	440	2	US-09-684-855-129	Sequence 129, App
34	125	5.5	440	2	US-09-684-855-152	Sequence 152, App
35	125	5.5	440	2	US-09-488-265B-7	Sequence 7, Appli
36	125	5.5	467	2	US-09-684-855-165	Sequence 165, App
37	124	5.5	467	2	US-09-488-265B-33	Sequence 33, App
38	118	5.2	440	2	US-09-684-855-108	Sequence 108, App
39	118	5.2	440	2	US-09-684-855-130	Sequence 130, App
40	118	5.2	440	2	US-09-684-855-153	Sequence 153, App
41	118	5.2	440	2	US-09-488-265B-8	Sequence 8, Appli
42	118	5.2	465	2	US-09-044-718-80	Sequence 80, Appl
43	118	5.2	465	2	US-10-062-848-80	Sequence 80, Appl
44	117	5.2	465	2	US-09-044-718-79	Sequence 79, Appl
45	117	5.2	465	2	US-10-062-848-79	Sequence 79, Appl

ALIGNMENTS

RESULT 1
US-09-866-379-10
; Sequence 10, Application US/09866379
; Patent No. 6855365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: O'DONOGHUE, Eileen
; APPLICANT: GARRETT, James
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 10
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Modified phytase
US-09-866-379-10

Query Match   98.7%; Score 2228; DB 2; Length 430;
Best Local Similarity   99.5%; Pred. No. 2.3e-227;
Matches 430; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAITADVDERTRKTGE 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAITADVDERTRKTGE 120

Qy   121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180

```
RESULT 2
US-09-866-379-8
; Sequence 8, Application US/09866379
; Patent No. 6855365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 8
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-866-379-8

Query Match             96.6%;  Score 2182;  DB 2;  Length 432;
Best Local Similarity   98.1%;  Pred. No. 1.8e-222;
Matches   424;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISVSRHGVRAPTKATQLMQDVTPDAWP   60
         ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP    60

Qy   61  TWPVKLGWLTPRGGELIAYLGHYWRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE   120
         |||||||.|||||||||||||||||||||||| |.|||||||||||||||||||||||||
Db   61  TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE   120

Qy  121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH   180
         |.|||||||||||||||||||||.|||||||||||||||||||||||| |||||||||||
Db  121  APAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILE-AGGSIADFTGH   179

Qy  181  YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT   240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  180  YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT   239

Qy  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA   300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  240  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA   298

Qy  301  LTPHPPQOKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW   360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  299  LTPHPPQOKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW   358

Qy  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ   420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  359  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ   418

Qy  421  IVNEARIPACSL   432
         ||||||||||||
Db  419  IVNEARIPACSL   430
```

```
RESULT 3
US-09-715-477-1
; Sequence 1, Application US/09715477
; Patent No. 6841370
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
; FILE REFERENCE: 19603/4031
; CURRENT APPLICATION NUMBER: US/09/715,477
; CURRENT FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-715-477-1

Query Match             96.4%;  Score 2177;  DB 2;  Length 432;
Best Local Similarity   97.9%;  Pred. No. 6.1e-222;
Matches   423;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISVSRHGVRAPTKATQLMQDVTPDAWP   60
         ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP    60

Qy   61  TWPVKLGWLTPRGGELIAYLGHYWRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE   120
         |||||||.|||||||||||||||||||||||| |.|||||||||||||||||||||||||
Db   61  TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE   120

Qy  121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH   180
         |.|||||||||||||||||||||.|||||||||||||||||||||||||||||||||||
Db  121  APAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH   180

Qy  181  YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT   240
         |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db  181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT   240

Qy  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA   300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA   300

Qy  301  LTPHPPQOKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW   360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  LTPHPPQOKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW   360

Qy  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ   420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ   420

Qy  421  IVNEARIPACSL   432
         ||||||||||||
Db  421  IVNEARIPACSL   432
```

Qy   421 IVNEARIPACSL 432
           ||||||||||||
Db   421 IVNEARIPACSL 432

RESULT 4
US-09-259-214-2
; Sequence 2, Application US/09259214A
; Patent No. 6110719
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: DIVER1370-1
; CURRENT APPLICATION NUMBER: US/09/259,214A
; CURRENT FILING DATE: 1999-03-01
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-259-214-2

Query Match   96.1%;   Score 2171;   DB 2;   Length 440;
Best Local Similarity   97.7%;   Pred. No. 2.7e-221;
Matches 422;   Conservative   0;   Mismatches 10;   Indels   0;   Gaps   0;

Qy   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy 121 AFAAGLAPDCAITVHTQADTSSPPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
          |||||||||||||||||||||| ||||||||||||||||||||||| |||||||||||
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy 181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
          |||||||||||||||||||||||||| |||||||||||||||||| ||||||||||||
Db 181 RQTAFRELERVLNFPQSNLCLKREKQDSCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy 241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
          |||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db 241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy 301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLLPGQPDNTPPGGELVFERW 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLLPGQPDNTPPGGELVFERW 360

Qy 361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy 421 IVNEARIPACSL 432
          ||||||||||||
Db 421 IVNEARIPACSL 432

RESULT 5
US-09-318-528-2
; Sequence 2, Application US/09318528
; Patent No. 6183740
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/318,528
; CURRENT FILING DATE: 1999-05-25

; EARLIER APPLICATION NUMBER: 09/291,931
; EARLIER FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-318-528-2

Query Match   96.1%;   Score 2171;   DB 2;   Length 440;
Best Local Similarity   97.7%;   Pred. No. 2.7e-221;
Matches 422;   Conservative   0;   Mismatches 10;   Indels   0;   Gaps   0;

Qy   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy 121 AFAAGLAPDCAITVHTQADTSSPPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
          |||||||||||||||||||||| ||||||||||||||||||||||| |||||||||||
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy 181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
          |||||||||||||||||||||||||| |||||||||||||||||| ||||||||||||
Db 181 RQTAFRELERVLNFPQSNLCLKREKQDSCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy 241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
          |||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db 241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy 301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLLPGQPDNTPPGGELVFERW 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLLPGQPDNTPPGGELVFERW 360

Qy 361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy 421 IVNEARIPACSL 432
          ||||||||||||
Db 421 IVNEARIPACSL 432

RESULT 6
US-09-291-931-2
; Sequence 2, Application US/09291931A
; Patent No. 6190697
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/291,931A
; CURRENT FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; SOFTWARE: FastSEQ for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-291-931-2

RESULT 7
US-09-580-515-2
; Sequence 2, Application US/09580515
; Patent No. 6720014
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/580,515
; CURRENT FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; SEQ ID NO 2
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; TYPE: PRT
; LENGTH: 440
; ORGANISM: Escherichia coli
;
US-09-580-515-2

Query Match   96.1%; Score 2171; DB 2; Length 440;
Best Local Similarity   97.7%; Pred. No. 2.7e-221;
Matches 422; Conservative   0; Mismatches 10; Indels   0; Gaps   0;

Qy   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGELTPRGGELIAYLGHYWQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGH 120

Qy   121 APAAGLAPDCAITVHIQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy   181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy   241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVPERW 360
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVPERW 360

Qy   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy   421 IVNEARIPACSL 432
               ||||||||||||
Db   421 IVNEARIPACSL 432

RESULT 8
US-09-866-379-2
; Sequence 2, Application US/09866379
; Patent No. 6855365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; SEQ ID NO 2
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; TYPE: PRT
; LENGTH: 440
; ORGANISM: Escherichia coli
;
US-09-866-379-2

Query Match   96.1%; Score 2171; DB 2; Length 440;
Best Local Similarity   97.7%; Pred. No. 2.7e-221;
Matches 422; Conservative   0; Mismatches 10; Indels   0; Gaps   0;

Qy   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGELTPRGGELIAYLGHYWQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGH 120

RESULT 9

US-09-715-477-3

```
; Sequence 3, Application US/09715477
; Patent No. 6841370
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
; FILE REFERENCE: 19603/4031
; CURRENT APPLICATION NUMBER: US/09/715,477
; CURRENT FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
;
; SEQ ID NO 3
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-715-477-3
```

Query Match   95.5%;   Score 2157; DB 2; Length 432;
Best Local Similarity   97.2%; Pred. No. 8e-220;
Matches 420; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

```
Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
        ||||||||||||||||||||||| |||||||||| || |||| |||||||||||||||||
Db   61 TWPVKLGELTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQPGGVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180

Qy  181 YQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
        |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db  181 RQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy  421 IVNEARIPACSL 432
        ||||||||||||
Db  421 IVNEARIPACSL 432
```

RESULT 10

US-10-266-041A-1

```
; Sequence 1, Application US/10266041A
; Patent No. 6974690
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: PHOSPHATASESS WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2798
; CURRENT APPLICATION NUMBER: US/10/266,041A
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: 60/127,032
; PRIOR FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 09/540,149
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
;
; SEQ ID NO 1
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-266-041A-1
```

Query Match   95.1%; Score 2147; DB 2; Length 432;
Best Local Similarity   96.8%; Pred. No. 9.3e-219;
Matches 418; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

```
Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
        ||||||||||||||||||||||| |||||||||| || |||| |||||||||||||||||
Db   61 TWPVKLGELTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQPGGVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180

Qy  181 YQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
        ||||||||||||||| |||||||||||||||||||||||||||||| |||||||||||||
Db  181 RQTAFRELERVLNFSQLNLCLNREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy  421 IVNEARIPACSL 432
        ||||||||||||
Db  421 IVNEARIPACSL 432
```

RESULT 11

US-09-540-149A-1

```
; Sequence 1, Application US/09540149A
; Patent No. 6511699
; GENERAL INFORMATION:
```

APPLICANT: Lei, Xingen
TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
FILE REFERENCE: 19603/2791
CURRENT APPLICATION NUMBER: US/09/540,149A
CURRENT FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 60/127,032
PRIOR FILING DATE: 1999-03-31
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 433
TYPE: PRT
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: UNSURE
LOCATION: (433)
OTHER INFORMATION: Xaa at position 433 in this sequence is unknown
US-09-540-149A-1

Query Match   95.1%; Score 2147; DB 2; Length 433;
Best Local Similarity 96.8%; Pred. No. 9.3e-219;
Matches 418; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

```
Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
        ||||||||||||||||||||||||||||||||||  |||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYWRQRLVADGLLAKKGCPQPGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
        |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy  181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
        |||||||||||||| | |||| |||||||||||||||||||||||| |||||||||||||
Db  181 RQTAFRELERVLNFSQLNLCLNREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300
        |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFLLQRTPEVARSRATPLLDLIMAA 300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy  421 IVNEARIPACSL 432
        ||||||||||||
Db  421 IVNEARIPACSL 432
```

RESULT 12
US-08-910-798-2
Sequence 2, Application US/08910798
Patent No. 5876997
GENERAL INFORMATION:
APPLICANT: KRETZ
TITLE OF INVENTION: NOVEL PHYTASE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910,798
FILING DATE: August 13,1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: HAILE, PH.D., LISA A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 09010/029001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-910-798-2

Query Match   91.4%; Score 2063.5; DB 1; Length 423;
Best Local Similarity 93.8%; Pred. No. 6.5e-210;
Matches 405; Conservative 0; Mismatches 10; Indels 17; Gaps 1;

```
Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
        ||||||||||||||||||||||||||||||||||  |||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYWRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTG- 119

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
                       |||||||||||||||||||||||||||||||| ||||||||||||||
Db  120 ---------------QADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 163

Qy  181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
        |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db  164 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 223

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300
        |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db  224 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFLLQRTPEVARSRATPLLDLIMAA 283

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  284 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 343

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  344 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 403

Qy  421 IVNEARIPACSL 432
        ||||||||||||
Db  404 IVNEARIPACSL 415
```

RESULT 13
US-09-489-039A-7512
Sequence 7512, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342

```
; SEQ ID NO 7512
;   LENGTH: 421
;   TYPE: PRT
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7512

  Query Match             24.0%;     Score 543;  DB 2;  Length 421;
  Best Local Similarity   32.6%;     Pred. No. 1.5e-48;
  Matches 136;  Conservative 70;  Mismatches 181;  Indels 30;  Gaps 10;

Qy   19  SARAQSE---PE-LKLESVVISRHGVRAP-TKATQLMQDVTPDAWPTWPVKLGELTPR   72
              ::: :   : :  |||: |||||:||||  |||  |||||||||||||||||||:: |
Db   23  SAGAQAQDKAAPEGYQLQQVTLMSRHNLRAPLANNGSVLEQSTAKAWPQWDVPCGQLTTK  82

Qy   73  GGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTKTGERAPAAGLAPDCAI  132
         ||  : |.|||| : || :.|||  | || | |||  : :||||||| : :  : ||||||
Db   83  GGVLEVYMGHYMREWLAQQKLVTSGECPPENAVAYANSLQRTVATAQPFITGAPPGCGI  142

Qy  133  TVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAI-LERAGGSIADFTGHYQTAFRELERV  191
Db  143  PVHHQPQMGTMDPTFNPVITDDSPAFREKALQAMEKERQGMQLTE------SYKLLETM   195

Qy  192  LNFPQOSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLTEIFLLQQAQGM  251
Db  196  IDYRNSPSC----KEKVCSLSEGKDTFSAGYQOEPGVSGPLKVGNSLVDAFTLQYEGF    251

Qy  252  P--BPGWGRITDSHQWNTLLSLLHNAQFDLLQRTPEVARSRATPLLDLIKTALTPHPPQKQ  309
Db  252  PKDQVAWGEIASDKQWRVLSKLNGYQDSLFTSVAVAQNVAKPLVKYIDNALVGEBGASK-  310

Qy  310  AYGVTLPTSVLFIAGHDTNLANLGGALEIN-WTLPGQPDNTPGGELVFERWRRLSDNSQ   368
Db  311  -----AKVTLLVGHDSNIASLLTALDPKPYQLPGQYERTPIGGKLLFQRWMDSAGNRD   363

Qy  369  WIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAAQGMCSLAGFTQIVNEA   425
Db  364  LMKIEYVYQSTEQLRNADALITLQAPPQRVTLALNGCPV-DDQGFCPLETFKMVINEA   419
```

```
RESULT 14
US-09-489-039A-13501
; Sequence 13501, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;   APPLICANT: Gary Breton et. al
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;   TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;   FILE REFERENCE: 2709.2004001
;   CURRENT APPLICATION NUMBER: US/09/489,039A
;   CURRENT FILING DATE: 2000-01-27
;   PRIOR APPLICATION NUMBER: US 60/117,747
;   PRIOR FILING DATE: 1999-01-29
;   NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13501
;   LENGTH: 522
;   TYPE: PRT
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13501

  Query Match             16.9%;     Score 381;  DB 2;  Length 522;
  Best Local Similarity   26.3%;     Pred. No. 3.3e-31;
  Matches 117;  Conservative 73;  Mismatches 209;  Indels 46;  Gaps 15;

Qy    2  KAILIPFLSLIIPLTP-QSAFAQSEPELKLESVVISRHGVRAPTKAT-QLMQDVTPDAW   59
Db  107  QGLLRLFIACALPLLALQSAAA---ADWQLEKVELSRHGIRPPTAGWREATEAATGRPW  163

Qy   60  PTWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTKTG  119
Db  164  TBWTTHDGELTGHGYAAVVNKGRAEGQHYRQLGLL-QAGCPTAESIVRASPLQRTRATA  222

Qy  120  EAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTG  179
```

```
Db  223  QALVDGAFPGCGVAIHYV--SGDADPLFQTDKFAATQTDPARQLAAVKEKAG---DLAQ  276
Qy  277  RRQA---LAPTIQLLKQAVC--QADKPCPIFDT-PWQVEQSKSGKTIISGLSVWANM    327

Db  240  TBIFLLQOAQGMP--EPGWGRITDSHQWNTLLSLLHNAQFDLLQRTPEVARSRATPLLDLI  297
Qy  328  VETLRLGWSENLPLSQLAWGKIITQARQITPALLPLLTENVDLTNDVLYTAQKRGSVLLNAM  387

Db  180  HYQTAFRELERVLNFPQOSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASML  239
Qy  388  LDGVKPEANPNVRW------LLIVAHDTNIAMVRTLMNFSWQLPGYSRGNIPGGSSLV    439

Db  357  FERRRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLT-----LAGCEERNAAQ  411
Qy  440  LERWRNAKSGERYLRVYFQOAQGLDDIR------RLQTPDAQHPMLRQBWHQPGCRQTDVGT  494

Db  412  MC-----SLAGFTQIVNEARIPACSL  432
Qy  495  LCPFQOAAITALGQRIDRSSAPAVAM  519
```

```
RESULT 15
US-09-044-718-3
; Sequence 3, Application US/09044718
; Patent No. 6391605
; GENERAL INFORMATION:
;   APPLICANT: KOSTREWA, Dirk
;   APPLICANT: PASAMONTES, Luis
;   APPLICANT: TOMSCHY, Andrea
;   APPLICANT: van LOON, Adolphus
;   APPLICANT: VOGEL, Kurt
;   APPLICANT: WYSS, Markus
;   TITLE OF INVENTION: MODIFIED PHYTASES
;   FILE REFERENCE: Modified Phytases
;   CURRENT APPLICATION NUMBER: US/09/044,718
;   CURRENT FILING DATE: 1998-03-19
;   PRIOR APPLICATION NUMBER: EP 97810175.6
;   PRIOR FILING DATE: 1997-03-25
;   NUMBER OF SEQ ID NOS: 82
;   SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 439
;   TYPE: PRT
;   ORGANISM: Aspergillus fumigatus
US-09-044-718-3

  Query Match             5.6%;      Score 126;  DB 2;  Length 439;
  Best Local Similarity   20.9%;     Pred. No. 0.00028;
  Matches 102;  Conservative 69;  Mismatches 166;  Indels 150;  Gaps 26;

Qy    7  PFLSLILIPLTPQSAFAQSEPELKLESVVISRHGVRAPT----KATQLMQDVTPDA--   58
Db   26  PFFSLEDELSVSSKLPK---DCRITLVQVLSRIGARYPFTSSKSKKYKGLVTAIQANATDP  82

Qy   59  -----WPTWPVKLG--ELTPRGGELIAYLG--HYWRQRLVADGLLPKCGCPQSGQVAII  108
Db   83  KGKPAFLKTYNTLGADDLTPFGEQOLVNSGIKFYQRYKALARSVVP---------PIR    132

Qy  109  ADVDERTKTGERAPAAGL-------------APDCAITVHTQADTSSPDPLFNPLKTGV   154
Db  133  ASGSSDRVIASGEKPIEGFCQOAKLADPGATNRAAPAISVII-PESETFN----NTLDHGV  186

Qy  155  C-QLDNANVTDAILERAGGSIADFTGHYQTAFRELERVLNFPQOSNLCLKREKQDESCSLT  213
Db  187  CTKFEASQLGDEV------AANFTALFAPDIR------                        218

Qy  214  QALPSELKVSADCVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSL--  271
Db  219  --LPGVTLTDBDVVSLMDMCSFDTV-------                              258

Qy  272  HN--AQFDLLQR------------TPE-VARSRATPLLDLIKT--ALTPHP          305
```

```
Db  259 HNEWKKYNYLQSLGKYYGYGAGNPLGPAQGIGFTNELIARLTRSPVQDHTSTNSTLVSNP 318

Qy  306 PQKQAYGVTLPTSVLPIA--GHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERWRRL 363

Db  319 ------ATPPLNATMYVDFSHDNSMVSIFPAL-------GLYNGTEPLSRTSVESAKEL 364

Qy  364 SD-NSQWI---QVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFT 419

Db  365 DGYSASWVVPFGARAYFPETMQCKSEKEPL-VRALINDRVVPLHGCDV-DKLGRCKLNDFV 422

Qy  420 QIVNEAR 426

Db  423 KGLSWAR 429
```

Search completed: June 13, 2006, 10:30:46
Job time : 36.6697 secs

OM protein - protein search, using sw model

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

Run on: June 13, 2006, 10:46:00 ; Search time 9.90826 Seconds
                                               (without alignments)
                                               551.244 Million cell updates/sec

Title: US-10-601-319-10
Perfect score: 2258
Sequence: 1 MKAILIPFLSLLIPLTPQSA..............CSLAGFTQIVNEARIPACSL 432

Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched: 64916 seqs, 12643201 residues

Total number of hits satisfying chosen parameters: 64916

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published Applications AA_New:*
                1: /EMC_Celerra_SIDS3/pcodata/1/pubpaa/US09_NEW_PUB.pep:*
                2: /EMC_Celerra_SIDS3/pcodata/1/pubpaa/US06_NEW_PUB.pep:*
                3: /EMC_Celerra_SIDS3/pcodata/1/pubpaa/US07_NEW_PUB.pep:*
                4: /EMC_Celerra_SIDS3/pcodata/1/pubpaa/US08_NEW_PUB.pep:*
                5: /EMC_Celerra_SIDS3/pcodata/1/pubpaa/PCT_NEW_PUB.pep:*
                6: /EMC_Celerra_SIDS3/pcodata/1/pubpaa/US10_NEW_PUB.pep:*
                7: /EMC_Celerra_SIDS3/pcodata/1/pubpaa/US11_NEW_PUB.pep:*
                8: /EMC_Celerra_SIDS3/pcodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	101.5	4.5	635	6	US-10-511-937-2424	Sequence 2424, Ap
2	99	4.4	384	7	US-11-293-697-2758	Sequence 2758, Ap
3	93.5	4.1	765	7	US-11-317-329-3	Sequence 3, Appli
4	93.5	4.1	765	7	US-11-317-329-34	Sequence 34, Appl
5	93.5	4.1	765	7	US-11-317-329-35	Sequence 35, Appl
6	93.5	4.1	765	7	US-11-317-329-36	Sequence 36, Appl
7	93.5	4.1	765	7	US-11-317-329-37	Sequence 37, Appl
8	93.5	4.1	765	7	US-11-317-329-38	Sequence 38, Appl
9	93.5	4.1	765	7	US-11-317-329-39	Sequence 39, Appl
10	93.5	4.1	765	7	US-11-317-329-40	Sequence 40, Appl
11	93.5	4.1	765	7	US-11-317-329-41	Sequence 41, Appl
12	93.5	4.1	765	7	US-11-317-329-42	Sequence 42, Appl
13	93.5	4.1	765	7	US-11-317-329-43	Sequence 43, Appl
14	93.5	4.1	765	7	US-11-317-329-44	Sequence 44, Appl
15	90.5	4.0	5738	6	US-10-505-928-150	Sequence 150, App
16	93.5	3.8	1127	6	US-10-527-411-40	Sequence 40, Appl
17	85.5	3.8	1127	6	US-10-527-411-50	Sequence 50, Appl
18	85.5	3.8	1127	6	US-10-527-411-54	Sequence 54, Appl
19	85.5	3.8	1127	6	US-10-527-411-58	Sequence 58, Appl
20	85.5	3.8	1129	6	US-10-527-411-42	Sequence 42, Appl
21	85.5	3.8	1129	6	US-10-527-411-48	Sequence 48, Appl
22	85.5	3.8	1129	6	US-10-527-411-52	Sequence 52, Appl
23	85.5	3.8	1129	6	US-10-527-411-56	Sequence 56, Appl
24	85.5	3.8	1130	6	US-10-527-411-44	Sequence 44, Appl
25	85.5	3.8	1130	6	US-10-527-411-139	Sequence 139, App
26	85.5	3.8	1132	6	US-10-527-411-46	Sequence 46, Appl
27	85.5	3.8	4391	7	US-11-183-325-56	Sequence 56, Appl
28	85	3.8	457	7	US-10-953-349-11735	Sequence 11735, A
29	85	3.8	460	6	US-10-953-349-34904	Sequence 34904, A
30	85	3.8	464	6	US-10-953-349-34903	Sequence 34903, A
31	85	3.8	472	6	US-10-953-349-34902	Sequence 34902, A
32	85	3.8	485	6	US-10-953-349-32471	Sequence 32471, A
33	85	3.8	504	6	US-10-953-349-32470	Sequence 32470, A
34	85	3.8	528	6	US-10-953-349-32469	Sequence 32469, A
35	85	3.8	537	6	US-10-953-349-11734	Sequence 11734, A
36	85	3.8	545	6	US-10-953-349-11733	Sequence 11733, A
37	84	3.7	847	6	US-10-953-349-300	Sequence 300, App
38	83	3.7	1730	7	US-10-505-928-284	Sequence 284, App
39	82	3.6	620	6	US-10-505-928-300	Sequence 300, App
40	81.5	3.6	725	7	US-11-293-697-3258	Sequence 3258, A
41	80.5	3.6	351	6	US-11-270-040-6	Sequence 6, Appli
42	80	3.5	760	6	US-10-505-928-50	Sequence 50, App
43	80	3.5	9535	6	US-10-471-571A-4496	Sequence 4496, A
44	79.5	3.5	800	6	US-10-953-349-33871	Sequence 33871, A
45	79.5	3.5	809	6	US-10-953-349-33870	Sequence 33870, A

ALIGNMENTS

RESULT 1
US-10-511-937-2424
; Sequence 2424, Application US/10511937
; Publication No. US20060088836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 50661200104
; CURRENT APPLICATION NUMBER: US 10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2424
; LENGTH: 635
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2424

  Query Match   4.5%; Score 101.5; DB 6; Length 635;
  Best Local Similarity   20.4%; Pred. No. 0.23;
  Matches 115; Conservative 60; Mismatches 195; Indels 193; Gaps 27;

Qy   2 KAILIPFLSLLIPLTPQSAFAQSEP-ELKL--ESVVIVS--------RHGVRAPTKAT-- 48
              ::::: ::: :   : ::: :: : :::   : ::: :: ::::
Db   123 RVLFVDSVGLPAPPSIIKAMGGSQPGELQISWEEPAPEISDFLRYELRYGPRDPKNSTGP 182

Qy   49 ---QLMQ--------------------DVTPDAWPTWPVKLG--------------ELTPRGGE-L 76
                 :::   :: : ::: ::::   : :: :
Db   183 TVIQLIATETCCPALQRPHSASALDQSPCAAQFTMPWQDGPKQTSPSREASALTABGGSCL 242

Qy   77 IAYL---GHYWRQ--RLVADGLLPKCCGCPQSGQVAIIADVDERTRKTGEAPAAGLAPDCAI 132
              ::::   ::::   : : :: : : ::   : :: ::: : :: : :
Db   183 ISGLQPGNSYWLQLRSEPDG-ISLGGSWGSWSLPVTVDL-------PGDAVALGL------ 289

```
Qy  133 TVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGHYQTAFRELERVL 192
Db  290 ----QCFT------------LDLKNVTCQWQQQDHASSQGFFYHSRARCCPRDR-- 327
Qy  193 NFPQSNLCLKREK----------QDBSCSLTQALPSELKVSADCVSLTGAVSLASMLTEIFL 244
Db  328 -YPIWENCEEEEKTNPGLQTPQFSRCHFKSRNDSIHILVEVTTAPGTVH--SYLGSPFW 384
Qy  245 LQQAQGMPEPG--WGRITDSH------------------------QWNTLLSLHNA 274
Db  385 IHQAVRLPTPNLHWREISSGHLELEWQHPSSWAAQETCYQLRYTGEHGQDWKVLEPPLGA 444
Qy  275 QFDLLQRTPEVARSRATPLLDLIKTALTPHPP--------------QKQAYGVTLPTSVLFI 322
Db  445 RGGTLELRP---RSRYRLQLRARLNGPTYQGFWSSWSDPTRVETATETAWISLVTALHLV 501
Qy  323 AGHDTNLANLGGALELNWTLPCQ--------------------PD------NTPPGGBLIV 356
Db  502 LG----LSAVLGLILLRWQPPAHYRRLRHALWPSLPDLHRVLGQYLRDTAALSPFKA--- 554
Qy  357 FERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEV------------KLTLAGCE 405
Db  555 -----TVSDTCEEVEPSLL-EILPKSSERTLPLCSSQAQMDYRRLQPSCLGTMPLSVCP 608
Qy  406 ERNAQGMCSLAGFTQIVNEARIP 428
Db  609 PMABSGSCCT---THIANHSYLP 628
```

RESULT 2
US-11-293-697-2758
Sequence 2758, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2758
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-2758

Query Match 4.4%; Score 99; DB 7; Length 384;
Best Local Similarity 23.0%; Pred. No. 0.18;
Matches 65; Conservative 21; Mismatches 90; Indels 106; Gaps 13;

```
Qy    5 LIPFLSLLI----PLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAWP 60
Db   69 LLPELPTQISSSRPTRPRCKSSSRPHL-------ARGRRSPLKAAQ--QDRGLPAWG 117
Qy   61 TWPVKLGELTPRG--GELIAYLG-----------HYWRQRLVADGLLPKCGCPQSG 103
Db  118 VWLLSRAEVPRSALGSSRRGLGDKGPSRRAAEFTQARPTRKRRTPGGGRPEAGC---- 173
Qy  104 QVAIIADVDERRKTGEAFAAGLAPDCAITVHTQADTSSSPDLFNPLKTGVCQLDNAN-- 161
Db  174 -WLENRRQRRRKAG-AERAGAGP---TLTPPLPAGSPDP-SPAATAARPQPEQRPHL 224
Qy  162 ------VTDAILERAGGSIADFTGHYQTAFRELERVLNFPQSNLCLKREKQDESCSLTQ 214
Db  225 SRGQWRLLPRRPERAAGSAG-------------------------------SCSSRR 251
Qy  215 ALPSELKVSADCVSLTGAVSLASMLTEIFLLQQAQGMPEPGW 256
Db  252 ALPA----SGSGAAMTSWA----------------------APPRPAW 274
```

RESULT 3
US-11-317-329-3
Sequence 3, Application US/11317329
; Publication No. US20060105413A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
; FILE REFERENCE: UM-06967
; CURRENT APPLICATION NUMBER: US/11/317,329
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/128,174
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 765
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-317-329-3

Query Match 4.1%; Score 93.5; DB 7; Length 765;
Best Local Similarity 21.6%; Pred. No. 1.6;
Matches 88; Conservative 46; Mismatches 144; Indels 129; Gaps 18;

```
Qy    8 FLSLLIPLLTPQSAFAQSE------PELKLESVIVSRHGVRAPTKA----TQLMQDVTPDAWP 60
Db  278 FLDITIETDILLSLLQSRVAVPESKALLARKVSCKLSLRQPREVNEDISQELMDSDSGNYL 337
Qy   61 TWPVKLGE---LTPRGGELIAYLGHYW-RQRLVADGLLPKCGCPQSGCQVAII---ADVD 112
Db  338 KRALQLSDRKNLVPRDEELCIYENKVTPLQFLVAQGSVE-------QVRLLAHEVDVD 389
Qy  113 ERTRKTGEAFAAGLAP---------DCAITVHTQADTSSPD----PLFNPLKTGVCQ 156
Db  390 CQT------ASGYTPLLIAAQDQQPDLCALLLAHGADANRVDEDGWAPLHFAAAQNGDDR 442
Qy  157 -----LDNANVTDAILERAGGSIADFTGHYQTAFRELERVLNFPQSNLCLKREKQDESCS 211
Db  443 TARLLLDHGACVDA--QEREGWTPLHLAA--QNNFENVARLLVSRQADPNIRE------ 491
Qy  212 LTQALPSELKVSADCVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSL 271
Db  492 ------AEGKTPLHVAAYFGHVSLVKLLT----------------SQGAELDAQQRNLRTPL 531
Qy  272 HNAQFDLLQRTPEVARSRATPLLDLIKTALTPHPPQKQAYG-------------- 312
Db  532 HLA---------VERGKVRAIQHLLKSGAVPDDALDQSGYGPLHTAAARGKYLICKMLLR 581
Qy  313 ----VTLPT---SVLFIAGHDTNL---------ANLGGALRLNWT 341
Db  582 YGASLELPTHQGWTPLHAAYKGHLEIIHILLAESHANMGALGAVNWT 628
```

RESULT 4
US-11-317-329-34
Sequence 34, Application US/11317329
; Publication No. US20060105413A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
; FILE REFERENCE: UM-06967
; CURRENT APPLICATION NUMBER: US/11/317,329
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/128,174
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 34
; LENGTH: 765
; TYPE: PRT

```
; ORGANISM: Homo sapiens
US-11-317-329-34

Query Match    4.1%; Score 93.5; DB 7; Length 765;
Best Local Similarity 21.6%; Pred. No. 1.6;
Matches 88; Conservative 46; Mismatches 144; Indels 129; Gaps 18;

Qy    8 FISLLIPLTPQSAFAQSE---PELKLESVVIVSRHGVRAPTKA----TQLMQDVTPDAWP 60
          ::|  :    |||   ||:  :::  ||    :|::::
Db  278 FIDITIETDILLSLLQSRVAVPESKALARKVSCKLSLRQPREVNEDISQELMDSDSGNYL 337

Qy   61 TWPVKLGE---LTPRGGELIAYLGHYW-RQRLVADGLLPKCGCPQSGQVAII----ADVD 112
          ::|   :::  ||  |
Db  338 KRALQLSDRKNLVPRDEELCIYENKVTPLQFLVAQGSVE------QVRLLLAHEVDVD 389

Qy  113 ERTRKTGEAFAAGLAP-------------DCAITVHTQADTSSPD------PLFNPLKTGVCQ 156
            :|||                      |||              |||
Db  390 CQT-----------ASGYTPLLIAAQDQQPDLCALLLAHGADANRVDEDGWAPLHFAAQNGDDR 442

Qy  157 -----LDNANVTDAILERAGGSIADFTGHYQTAFRELERVLNFPQSNLCLKREKQDESCS 211
             |:|  |||
Db  443 TARLLLDHGACVDA--QEREGWTPLHLAA--QNNFENVARLLVSRQADPNLRE------- 491

Qy  212 LTQALPSELKVSADCVSLTGAVSLASMLTEIFILQQAQGMPEPGWGRITDSHQWNTLLSL 271
           |:|  ::|          |||
Db  492 -------AEGKTPLHVAAYPGHVSLVKLLT--------SQGAELDAQQRNLRTPL 531

Qy  272 HNAQPDLLQRTPEVARSRATPLLDLIKTALTPHPPQKQAYG------------------- 312
             :|::  |||
Db  532 HLA------VERGKVRAIQHLLKSGAVPDALDQSGYPLHTAAARGKYLICKMLLR 581

Qy  313 ----VTLPT---SVLFIAGHDTNL------------------ANLGGALELNWT 341
            |||||    :|||           |||
Db  582 YGASLELPTHQGWTPLHLAAYKGHLEIIHLLAESHANMGALGAVNWT 628
```

RESULT 5
US-11-317-329-35
; Sequence 35, Application US/11317329
; Publication No. US20060105413A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
; FILE REFERENCE: UM-06967
; CURRENT APPLICATION NUMBER: US/11/317,329
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/128,174
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 35
; LENGTH: 765
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-317-329-35

```
Query Match    4.1%; Score 93.5; DB 7; Length 765;
Best Local Similarity 21.6%; Pred. No. 1.6;
Matches 88; Conservative 46; Mismatches 144; Indels 129; Gaps 18;

Qy    8 FISLLIPLTPQSAFAQSE---PELKLESVVIVSRHGVRAPTKA----TQLMQDVTPDAWP 60
          ::|  :    |||   ||:  :::  ||    :|::::
Db  278 FIDITIETDILLSLLQSRVAVPESKALARKVSCKLSLRQPREVNEDISQELMDSDSGNYL 337

Qy   61 TWPVKLGE---LTPRGGELIAYLGHYW-RQRLVADGLLPKCGCPQSGQVAII----ADVD 112
Db  338 KRALQLSDRKNLVPRDEELCIYENKVTPLQFLVAQGSVE------QVRLLLAHEVDVD 389

Qy  113 ERTRKTGEAFAAGLAP-------------DCAITVHTQADTSSPD------PLFNPLKTGVCQ 156
Db  390 CQT-----------ASGYTPLLIAAQDQQPDLCALLLAHGADANRVDEDGWAPLHFAAQNGDDR 442

Qy  157 -----LDNANVTDAILERAGGSIADFTGHYQTAFRELERVLNFPQSNLCLKREKQDESCS 211
Db  443 TARLLLDHGACVDA--QEREGWTPLHLAA--QNNFENVARLLVSRQADPNLRE------- 491

Qy  212 LTQALPSELKVSADCVSLTGAVSLASMLTEIFILQQAQGMPEPGWGRITDSHQWNTLLSL 271
Db  492 -------AEGKTPLHVAAYPGHVSLVKLLT--------SQGAELDAQQRNLRTPL 531

Qy  272 HNAQPDLLQRTPEVARSRATPLLDLIKTALTPHPPQKQAYG------------------- 312
Db  532 HLA------VERGKVRAIQHLLKSGAVPDALDQSGYPLHTAAARGKYLICKMLLR 581

Qy  313 ----VTLPT---SVLFIAGHDTNL------------------ANLGGALELNWT 341
Db  582 YGASLELPTHQGWTPLHLAAYKGHLEIIHLLAESHANMGALGAVNWT 628
```

RESULT 6
US-11-317-329-36
; Sequence 36, Application US/11317329
; Publication No. US20060105413A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
; FILE REFERENCE: UM-06967
; CURRENT APPLICATION NUMBER: US/11/317,329
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/128,174
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 36
; LENGTH: 765
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-317-329-36

```
Query Match    4.1%; Score 93.5; DB 7; Length 765;
Best Local Similarity 21.6%; Pred. No. 1.6;
Matches 88; Conservative 46; Mismatches 144; Indels 129; Gaps 18;

Qy    8 FISLLIPLTPQSAFAQSE---PELKLESVVIVSRHGVRAPTKA----TQLMQDVTPDAWP 60
Db  278 FIDITIETDILLSLLQSRVAVPESKALARKVSCKLSLRQPREVNEDISQELMDSDSGNYL 337

Qy   61 TWPVKLGE---LTPRGGELIAYLGHYW-RQRLVADGLLPKCGCPQSGQVAII----ADVD 112
Db  338 KRALQLSDRKNLVPRDEELCIYENKVTPLQFLVAQGSVE------QVRLLLAHEVDVD 389

Qy  113 ERTRKTGEAFAAGLAP-------------DCAITVHTQADTSSPD------PLFNPLKTGVCQ 156
Db  390 CQT-----------ASGYTPLLIAAQDQQPDLCALLLAHGADANRVDEDGWAPLHFAAQNGDDR 442

Qy  157 -----LDNANVTDAILERAGGSIADFTGHYQTAFRELERVLNFPQSNLCLKREKQDESCS 211
Db  443 TARLLLDHGACVDA--QEREGWTPLHLAA--QNNFENVARLLVSRQADPNLRE------- 491

Qy  212 LTQALPSELKVSADCVSLTGAVSLASMLTEIFILQQAQGMPEPGWGRITDSHQWNTLLSL 271
Db  492 -------AEGKTPLHVAAYPGHVSLVKLLT--------SQGAELDAQQRNLRTPL 531

Qy  272 HNAQPDLLQRTPEVARSRATPLLDLIKTALTPHPPQKQAYG------------------- 312
Db  532 HLA------VERGKVRAIQHLLKSGAVPDALDQSGYPLHTAAARGKYLICKMLLR 581

Qy  313 ----VTLPT---SVLFIAGHDTNL------------------ANLGGALELNWT 341
Db  582 YGASLELPTHQGWTPLHLAAYKGHLEIIHLLAESHANMGALGAVNWT 628
```

RESULT 7
US-11-317-329-37

```
; Sequence 37, Application US/11317329
; Publication No. US20060105413A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
; FILE REFERENCE: UM-06967
; CURRENT APPLICATION NUMBER: US/11/317,329
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/128,174
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 37
; LENGTH: 765
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-317-329-37

Query Match    4.1%; Score 93.5; DB 7; Length 765;
Best Local Similarity   21.6%; Pred. No. 1.6;
Matches  88; Conservative  46; Mismatches 144; Indels 129; Gaps  18;

Qy      8 FLSLILPLTPQSAFAQSE-----PELKLESVVIVSRHGVRAPTKA-----TQLMQDVTPDAWP 60
          ||:|                   :||                        :  :
Db    278 FLDITIETDILLSLLQSRVAVPESKALARKVSCKLSLRQPREVNEDISQELMDSDSGNYL 337

Qy     61 TWPVKLGE---LTPRGGELIAYLGHYW-RQRLVADGLLPKCCGCPQSGQVAII-----ADVD 112
          :::|          ||              |||: ||                  ||
Db    338 KRALQLSDRKNLVPRDEELCIYENKVTPLQFLVAQGSVE------QVRLLAHEVDVD 389

Qy    113 ERTRKTGEAFAAGLAP--------------DCAITVHTQADTSSPD-----PLFNPLKTGVCQ 156
          ||                            |||                     |
Db    390 CQT------ASGYTPLLIAAQDQOPDLCALLLAHGADANRVDEDGWAPLHFAAQNGDDR 442

Qy    157 ----LDNANVTDAILERAGGSIADFTGHYQTAFRELERVLNFPQSNLCLKREKQDESCS 211
          |:|                   |:|                             |:|
Db    443 TARLLLDHGACVDA-QEREGWTPLHLAA--QNNFENVARLLVSRQADPNLRE------ 491

Qy    212 LTQALPSELKVSADCVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSL 271
          :|:                    |||||                          |
Db    492 ------ABGKTPLHVAAYFGHVSLVKLLT------SQGAELDAQORNLRTPL 531

Qy    272 HNAQFDLLQRTPEVARSRATPLLDLLIKTALTPHPPQKQAYG----- 312
          |||                   |:|
Db    532 HLA------VERGKVRAIQHLLKSGAVPDALDQSGYGPLHTAAARGKYLICKMLLR 581

Qy    313 ----VTLPT---SVLFIAGHDTNL----------ANLGGALELNWT 341
          :|||     :|:                      |||
Db    582 YGASLELPTHQGWTPLHLAAYKGHLEIIHLLAESHANMGALGAVNWT 628

RESULT 8
US-11-317-329-38
; Sequence 38, Application US/11317329
; Publication No. US20060105413A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
; FILE REFERENCE: UM-06967
; CURRENT APPLICATION NUMBER: US/11/317,329
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/128,174
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 38
; LENGTH: 765
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-317-329-38
```

```
Query Match    4.1%; Score 93.5; DB 7; Length 765;
Best Local Similarity   21.6%; Pred. No. 1.6;
Matches  88; Conservative  46; Mismatches 144; Indels 129; Gaps  18;

RESULT 9
US-11-317-329-39
; Sequence 39, Application US/11317329
; Publication No. US20060105413A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
; FILE REFERENCE: UM-06967
; CURRENT APPLICATION NUMBER: US/11/317,329
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/128,174
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 39
; LENGTH: 765
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-317-329-39

Db    582 YGASLELPTHQGWTPLHLAAYKGHLEIIHLLAESHANMGALGAVNWT 628
          :|||     :|:                      |||
Qy    313 ----VTLPT---SVLFIAGHDTNL----------ANLGGALELNWT 341

Db    532 HLA------VERGKVRAIQHLLKSGAVPDALDQSGYGPLHTAAARGKYLICKMLLR 581
          |||                   |:|
Qy    272 HNAQFDLLQRTPEVARSRATPLLDLLIKTALTPHPPQKQAYG----- 312

Db    492 ------ABGKTPLHVAAYFGHVSLVKLLT------SQGAELDAQORNLRTPL 531
          :|:                    |||||                          |
Qy    212 LTQALPSELKVSADCVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSL 271

Db    443 TARLLLDHGACVDA-QEREGWTPLHLAA--QNNFENVARLLVSRQADPNLRE------ 491
          |:|                   |:|                             |:|
Qy    157 ----LDNANVTDAILERAGGSIADFTGHYQTAFRELERVLNFPQSNLCLKREKQDESCS 211

Db    390 CQT------ASGYTPLLIAAQDQOPDLCALLLAHGADANRVDEDGWAPLHFAAQNGDDR 442
          ||                            |||                     |
Qy    113 ERTRKTGEAFAAGLAP--------------DCAITVHTQADTSSPD-----PLFNPLKTGVCQ 156

Db    338 KRALQLSDRKNLVPRDEELCIYENKVTPLQFLVAQGSVE------QVRLLAHEVDVD 389
          :::|          ||              |||: ||                  ||
Qy     61 TWPVKLGE---LTPRGGELIAYLGHYW-RQRLVADGLLPKCCGCPQSGQVAII-----ADVD 112

Db    278 FLDITIETDILLSLLQSRVAVPESKALARKVSCKLSLRQPREVNEDISQELMDSDSGNYL 337
          ||:|                   :||                        :  :
Qy      8 FLSLILPLTPQSAFAQSE-----PELKLESVVIVSRHGVRAPTKA-----TQLMQDVTPDAWP 60
```

RESULT 10
US-11-317-329-40
; Sequence 40, Application US/11317329
; Publication No. US20060105413A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
; FILE REFERENCE: UM-06967
; CURRENT APPLICATION NUMBER: US/11/317,329
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/128,174
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 40
; LENGTH: 765
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-317-329-40

Query Match   4.1%; Score 93.5; DB 7; Length 765;
Best Local Similarity   21.6%;
Matches   88; Conservative   46; Mismatches 144; Indels 129; Gaps   18;

```
Qy     8 FLSLILPLTPQSAPAQSE---PELKLESVVIVSRHGVRAPTKA----TQLMQDVTPDAWP 60
           ||| :                :  :   ||
Db   278 FLDITIETDILLSLLQSRVAVPESKALARKVSCKLSLRQPREVNEDISQELMDSDSGNYL 337

Qy    61 TWPVKLGE---LTPRGGELIAYLGHYW-RQRLVADGLLPKCGCPQSGQVAII----ADVD 112
           ||||            ::               |||                   |||
Db   338 KRALQLSDRKNLVPRDEELCIYENKVTPLQFLVAQGSVE-------QVRLLLAHEVDVD 389

Qy   113 ERTRKTGEAPAAGLAP----------DCAITVHTQADTSSSPD-----PLFNPLKTGVCQ 156
           :                         :                    ||
Db   390 CQT-----ASGYTPLLIAAQDQOPDLCALLLAHGADANRVDEDGWAPLHFAAQNGDDR 442

Qy   157 -----LDNANVTDAILERAGGSIADFTGHYQTAFRELERVLNFPQSNLCLKREKQDESCS 211
                :   :
Db   443 TARLLLDHGACVDA--QEREGWTPLHLAA--QNNFENVARLLVSRQADPNLRE------- 491

Qy   212 LTQALPSELKVSADCVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSL 271
           ||                |||  :
Db   492 ------AEGKTPLHVAAYFGHVSLVKLLT------------------SQGAELDAQQRNLRTPL 531

Qy   272 HNAQFDLLQRTPEVARSRATPLLDLIKTALTPHPPQKQAYG--------------------- 312
           |||                         :::
Db   532 HLA------------VERGKVRAIQHLLKSGAVPDALDQSGYGPLHTAAARGKYLICKMLLR 581

Qy   313 ----VTLPT----SVLPIAGHDTNL-------------------ANLGGALELNWT 341
               |||         ::                          |||
Db   582 YGASLELPTHQGWTPLHLAAYKGHLEIIHLLAABSHANMGALGAVNWT 628
```

RESULT 11
US-11-317-329-41
; Sequence 41, Application US/11317329
; Publication No. US20060105413A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
; FILE REFERENCE: UM-06967
; CURRENT APPLICATION NUMBER: US/11/317,329
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/128,174
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 41
; LENGTH: 765
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-317-329-41

Query Match   4.1%; Score 93.5; DB 7; Length 765;
Best Local Similarity   21.6%;
Matches   88; Conservative   46; Mismatches 144; Indels 129; Gaps   18;

```
Qy     8 FLSLILPLTPQSAPAQSE---PELKLESVVIVSRHGVRAPTKA----TQLMQDVTPDAWP 60
           ||| :                :  :   ||
Db   278 FLDITIETDILLSLLQSRVAVPESKALARKVSCKLSLRQPREVNEDISQELMDSDSGNYL 337

Qy    61 TWPVKLGE---LTPRGGELIAYLGHYW-RQRLVADGLLPKCGCPQSGQVAII----ADVD 112
           ||||            ::               |||                   |||
Db   338 KRALQLSDRKNLVPRDEELCIYENKVTPLQFLVAQGSVE-------QVRLLLAHEVDVD 389

Qy   113 ERTRKTGEAPAAGLAP----------DCAITVHTQADTSSSPD-----PLFNPLKTGVCQ 156
           :                         :                    ||
Db   390 CQT-----ASGYTPLLIAAQDQOPDLCALLLAHGADANRVDEDGWAPLHFAAQNGDDR 442

Qy   157 -----LDNANVTDAILERAGGSIADFTGHYQTAFRELERVLNFPQSNLCLKREKQDESCS 211
                :   :
Db   443 TARLLLDHGACVDA--QEREGWTPLHLAA--QNNFENVARLLVSRQADPNLRE------- 491

Qy   212 LTQALPSELKVSADCVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSL 271
           ||                |||  :
Db   492 ------AEGKTPLHVAAYFGHVSLVKLLT------------------SQGAELDAQQRNLRTPL 531

Qy   272 HNAQFDLLQRTPEVARSRATPLLDLIKTALTPHPPQKQAYG--------------------- 312
           |||                         :::
Db   532 HLA------------VERGKVRAIQHLLKSGAVPDALDQSGYGPLHTAAARGKYLICKMLLR 581

Qy   313 ----VTLPT----SVLPIAGHDTNL-------------------ANLGGALELNWT 341
               |||         ::                          |||
Db   582 YGASLELPTHQGWTPLHLAAYKGHLEIIHLLAABSHANMGALGAVNWT 628
```

RESULT 12
US-11-317-329-42
; Sequence 42, Application US/11317329
; Publication No. US20060105413A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
; FILE REFERENCE: UM-06967
; CURRENT APPLICATION NUMBER: US/11/317,329
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/128,174
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 42
; LENGTH: 765
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-317-329-42

Query Match   4.1%; Score 93.5; DB 7; Length 765;
Best Local Similarity   21.6%;
Matches   88; Conservative   46; Mismatches 144; Indels 129; Gaps   18;

RESULT 13
US-11-317-329-43
; Sequence 43, Application US/11317329
; Publication No. US20060105413A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
; FILE REFERENCE: UM-06967
; CURRENT APPLICATION NUMBER: US/11/317,329
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/128,174
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 43
; LENGTH: 765
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-317-329-43

Query Match   4.1%;   Score 93.5;   DB 7;   Length 765;
Best Local Similarity   21.6%;   Pred. No. 1.6;
Matches   88;   Conservative   46;   Mismatches   144;   Indels   129;   Gaps   18;

```
Qy    8 FLSLILPLTPQSAFAQSE-----PELKLESVVIVSRHGVRAPTKA----TQLMQDVTPDAWP   60
          :|:|  :                :|:|  |:  :   |:|                  :.|
Db  278 FLDITIETDILLSLLQSRVAVPESKALARKVSCKLSLRQPREVNEDISQELMDSDSGNYL  337

Qy   61 TWPVKLGE---LTPRGGELIAYLGHYW-RQRLIVADGLLPKCCGPQSGQVAII----ADVD  112
          ::|.|       |:|                               |||:   |||
Db  338 KRALQLSDRKNLVPRDEELCIYENKVTPLQFLVAQGSVE------QVRLLLAHEVDVD  389

Qy  113 ERTRKTGEAFAAGLAP-------DCAITVHTQADTSSPD-----PLFNPLKTGVCQ  156
          |||                    :||                   ||||
Db  390 CQT------ASGYTPLLIAAQDQOPDLCALLLAHGADANRVDEDGWAPLHFAAQNGDDR  442

Qy  157 -----LDNANVTDAILERAGGSIADFTGHYQTAFRELERVLNFPQSNICLKREKQDESCS  211
             |:||           ||:      :|                  |:
Db  443 TARLLLDHGACVDA--QEREGWTPLHLAA--QNNFENVARLLVSRQADPNLRE-------  491

Qy  212 LTQALPSELKVSADCVSLTGAVSLASMLTEIFLLQQOAOGMPEPGWGRITDSHQWNTLLSL  271
          :|:|           :|||
Db  492 -----AEGKTPLHVAAYFGHVSLVKLLT--------SQGAELDAQQRNLRTPL  531

Qy  272 HNAQFDLLQRTPEVARSRATPLLLDLIKTALTPHPPQKQAYG--------  312
             |||
Db  532 HLA-----VERGKVRAIQHLLKSGAVPDALDQSGYGPLHTAAARGKYLICKMLLR  581

Qy  313 ----VTLPT-----SVLFIAGHDTNL----------------ANLGGALELNWT  341
            :|||           |:|             :||   |||
Db  582 YGASLELPTHQGWTPLHLAAYKGHLEIIHLLAESHANMGALGAVNWT  628
```

RESULT 14
US-11-317-329-44
; Sequence 44, Application US/11317329
; Publication No. US20060105413A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
; FILE REFERENCE: UM-06967
; CURRENT APPLICATION NUMBER: US/11/317,329
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/128,174
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 44
; LENGTH: 765
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-317-329-44

Query Match   4.1%;   Score 93.5;   DB 7;   Length 765;
Best Local Similarity   21.6%;   Pred. No. 1.6;
Matches   88;   Conservative   46;   Mismatches   144;   Indels   129;   Gaps   18;

```
Qy    8 FLSLILPLTPQSAFAQSE-----PELKLESVVIVSRHGVRAPTKA----TQLMQDVTPDAWP   60
          :|:|  :                :|:|  |:  :   |:|                  :.|
Db  278 FLDITIETDILLSLLQSRVAVPESKALARKVSCKLSLRQPREVNEDISQELMDSDSGNYL  337

Qy   61 TWPVKLGE---LTPRGGELIAYLGHYW-RQRLIVADGLLPKCCGPQSGQVAII----ADVD  112
          ::|.|       |:|                               |||:   |||
Db  338 KRALQLSDRKNLVPRDEELCIYENKVTPLQFLVAQGSVE------QVRLLLAHEVDVD  389

Qy  113 ERTRKTGEAFAAGLAP-------DCAITVHTQADTSSPD-----PLFNPLKTGVCQ  156
          |||                    :||                   ||||
Db  390 CQT------ASGYTPLLIAAQDQOPDLCALLLAHGADANRVDEDGWAPLHFAAQNGDDR  442

Qy  157 -----LDNANVTDAILERAGGSIADFTGHYQTAFRELERVLNFPQSNICLKREKQDESCS  211
             |:||           ||:      :|                  |:
Db  443 TARLLLDHGACVDA--QEREGWTPLHLAA--QNNFENVARLLVSRQADPNLRE-------  491

Qy  212 LTQALPSELKVSADCVSLTGAVSLASMLTEIFLLQQOAOGMPEPGWGRITDSHQWNTLLSL  271
          :|:|           :|||
Db  492 -----AEGKTPLHVAAYFGHVSLVKLLT--------SQGAELDAQQRNLRTPL  531

Qy  272 HNAQFDLLQRTPEVARSRATPLLLDLIKTALTPHPPQKQAYG--------  312
             |||
Db  532 HLA-----VERGKVRAIQHLLKSGAVPDALDQSGYGPLHTAAARGKYLICKMLLR  581

Qy  313 ----VTLPT-----SVLFIAGHDTNL----------------ANLGGALELNWT  341
            :|||           |:|             :||   |||
Db  582 YGASLELPTHQGWTPLHLAAYKGHLEIIHLLAESHANMGALGAVNWT  628
```

RESULT 15
US-10-505-928-150
; Sequence 150, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178

```
;   CURRENT APPLICATION NUMBER: US/10/505,928
;   CURRENT FILING DATE:  2004-08-27
;   PRIOR APPLICATION NUMBER: US 60/363,019
;   PRIOR FILING DATE: 2002-03-07
;   NUMBER OF SEQ ID NOS: 866
;   SOFTWARE: Patentin 3.2
;
;   SEQ ID NO 150
;       LENGTH: 5738
;       TYPE: PRT
;       ORGANISM: Homo sapiens

US-10-505-928-150

  Query Match             4.0%;  Score 90.5;  DB 6;  Length 5738;
  Best Local Similarity   22.9%;  Pred. No. 53;
  Matches  110;  Conservative  35;  Mismatches 159;  Indels 177;  Gaps  29;

Qy      4 ILIPFLSLLIPLTPQSAP------------------AQSEPELKLESVVIVSRHGVR---APTKATQLMQD 53
          :|| |: || || || : |                  |: ||       : |   ||| :     : ||  ::
Db   2359 VVLPELQQPPPLRPRSPVPWHPQGTQTEPCEGCEHQGQVHRVGEBRWHGGPCRVCQCLHN 2418

Qy     54 VTPDAWPTWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDE 113
          :|| |  || : |||   ||| |   |  |  |          || |  |:
Db   2419 LTAHCSPYCP--LGS-CPGGWVLVEGTGB----------------------SCCHCALPGB------ 2454

Qy    114 RTRKTGEAFAAGLAPDCAITVHTQADTSSSPDPLFN-PLKTGVCQLDNANVTDAILERAGG 172
          |               : :||                  :    :          : : ||
Db   2455 ---NCTVQPMA------------------TPAAAPAPSPQIRFPLAT--------------YILPPSGG 2488

Qy    173 SIADFTGHYQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGA 232
          |                       :  |                     :     |:
Db   2489 SC------------------------RPLSSP-TPACL-------------SLLHPDP--------CYSPLGL 2515

Qy    233 VSLA------------------SMLTEIFLL-QQAQGMPEPGWGRITDSH-QWNTLLSLHNAQFDLL 279
                                 | |          : | :|   ||
Db   2516 AGLAEGSLHASSQQLEHPTQAALLGAPTQGPSPQGWHAGGDAYAKWHT--RPHYLQLDLL 2573

Qy    280 Q-----------------RTPEVARSRA---------------TPLLDLIKTALTPHPPQKQAYG---- 312
                             |                        |:   |   |   |||  |
Db   2574 QPRNLTGILVPETGSSNAYASSPSLQPSSNGLHWHDYRDLL-PGILPLPKVSPAQGRWGQ 2632

Qy    313 -VTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPG------GELVPER-WRR 362
               |   |       | |   :  |  |  | |           | |      |  |||
Db   2633 QPTMP-----FCGFHS-------------------LCPQGPSSVPBGHGLHSMLVEYLLFPRNWDD 2674

Qy    363 LSDNSQWIQVSLVPQTLQQMR---DKTPLSLNTPPGEVKLTLAGCEB-----RNAQGMCSLA 416
          |      |    :| |   | |  :       |:  :  | ||        |    |: |
Db   2675 L-DPAVW-----TFGRMVQARPVRVWPHDVHHSDVPLQVELLGCEPGVGVGLRCASGECVLR 2728

Qy    417 G 417
          |
Db   2729 G 2729
```

THIS PAGE LEFT BLANK

OM protein - protein search, using sw model

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

Run on: June 13, 2006, 10:22:24 ; Search time 25.2661 Seconds
                 (without alignments)
                 1645.118 Million cell updates/sec

Title: US-10-601-319-10
Perfect score: 2258
Sequence: 1 MKAILIPFLSLLIPLTPQSA.........CSLAGFTQIVNEARIPACSL 432

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   1: pir1:*
                 2: pir2:*
                 3: pir3:*
                 4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2182	96.6	432	2	B36733	acid phosphatase
2	2144	95.0	434	2	H90770	phosphoanhydride p
3	2144	95.0	444	2	D85633	hypothetical prote
4	959.5	42.5	441	2	AC0201	acid phosphatase (
5	560	24.8	413	2	F90773	periplasmic glucos
6	560	24.8	413	2	B85636	periplasmic glucos
7	554	24.5	413	2	JV0087	periplasmic glucos
8	547	24.2	413	2	B64839	glucose-1-phosphat
9	538	23.8	417	2	AG0632	glucose-1-phosphat
10	538	23.8	414	2	S25627	glucose-1-phosphat
11	536	23.7	414	2	E87316	periplasmic phosph
12	137	6.1	416	2	T16058	hypothetical prote
13	132	5.8	423	2	S06167	hypothetical prote
14	116.5	5.2	423	2	A33395	acid phosphatase (
15	112.5	5.0	438	2	S64682	acid phosphatase (
16	111.5	4.9	421	2	S41742	acid phosphatase (
17	109.5	4.8	344	2	B89130	protein P52E1.8 [1
18	109.5	4.8	457	2	A56925	paired box transcr
19	109.5	4.8	479	1	JN0890	paired box transcr
20	108	4.8	386	1	JH0610	acid phosphatase (
21	107.5	4.8	459	2	S52250	paired box transcr
22	107.5	4.8	479	1	JN0715	3-phytase (EC 3.1.
23	107	4.7	5126	2	S40450	ryanodine receptor
24	105	4.7	450	2	A54429	paired box transcr
25	104.5	4.6	397	2	C81716	hypothetical prote
26	103.5	4.6	537	2	S54770	secreted acid phos
27	103.5	4.6	888	2	T46726	secreted acid phos
28	103	4.6	1048	1	BVECSC	exonuclease (EC 3.
29	101.5	4.5	769	2	B87681	tyrosine kinase Di
30	101.5	4.5	356	2	P75594	hypothetical prote
31	101.5	4.5	635	2	A45266	MPL-P protein prec
32	101.5	4.5	789	2	AE0064	DNA-directed DNA p
33	101.5	4.5	1214	2	AG2897	conserved hypothet
34	100.5	4.5	1387	2	A97673	probable periplasm
35	100	4.4	425	2	D83186	hypothetical prote
36	99.5	4.4	971	2	T32883	hypothetical prote
37	99	4.4	234	2	F42696	thrombin (EC 3.4.2
38	99	4.4	381	2	F84341	acid phosphatase (
39	99	4.4	560	2	H96599	exonuclease SbcC [
40	98.5	4.4	969	2	AB0551	hypothetical prote
41	98	4.3	1034	2	T27997	protein F14J16.10
42	97.5	4.3	4056	2	JH0152	protein ZK792.1 [1
43	96.5	4.3	328	2	D88844	acid phosphatase (
44	96.5	4.3	374	2	A83572	pyridoxal phosphat
45	96.5	4.3	1047	2	C85535	ATP-dependent dsDN

ALIGNMENTS

RESULT 1
B36733

acid phosphatase (EC 3.1.3.2) precursor - Escherichia coli (strain K-12)
N;Alternate names: phosphoanhydride phosphohydrolase; periplasmic; phytase 6
C;Species: Escherichia coli
C;Date: 19-Apr-1991 #sequence revision 19-Apr-1991 #text change 09-Jul-2004
C;Accession: B36733; S18018; B64839; A26534; S17960; S33278
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: B64839
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-50, 'NAGCHPRRWANLAG', 65, 'T', 67-74, 'DV', 77-111, 'S' <TOU>
A;Cross-references: UNIPARC:UPI000016BDE2; GB:X05471; NID:g409925; PIDN:CAA29031.1; PID:g
A;Experimental source: strain K-12, substrain MG1655
Biochimie 69, 215-221, 1987
R;Touati, E.; Danchin, A.
A;Title: The structure of the promoter and amino terminal region of the pH 2.5 acid phos
A;Reference number: A26534; MUID:87271766; PMID:3038201
A;Accession: A26534
A;Molecule type: DNA
A;Residues: 1-50, 'NAGCHPRRWANLAG', 65, 'T', 67-74, 'DV', 77-111, 'S' <TOU>
A;Cross-references: UNIPARC:UPI000016DE2; GB:X05471; NID:g409925; PIDN:CAA29031.1; PID:g
J. Bacteriol. 172, 5497-5500, 1990
R;Dassa, J.; Marck, C.; Boquet, P. L.
A;Title: The complete nucleotide sequence of the Escherichia coli gene appA reveals sign
A;Reference number: A36733; MUID:90368616; PMID:2168385
A;Accession: B36733
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-432 <DAS>
A;Cross-references: UNIPROT:P07102; UNIPARC:UPI00004750A; GB:M58708; NID:g145283; PIDN
Biol. Chem. Hoppe-Seyler 372, 664-665, 1991
R;Greiner, R.; Jany, K. D.
A;Title: Characterization of a phytase from Escherichia coli.
A;Reference number: S18018
A;Accession: S18018
A;Molecule type: protein
A;Residues: 23-33 <GRE>
A;Cross-references: UNIPARC:UPI000001779815
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: B64839
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Cross-references: UNIPARC:UPI00000004750A; GB:AB000200; GB:U00096; NID:g2367111; PIDN:AA
A;Experimental source: strain K-12, substrain MG1655
A;Accession: A26534
A;Molecule type: DNA
A;Residues: 1-432 <BLAT>
A;Cross-references: UNIPARC:UPI0000004750A; GB:AB000200; GB:U00096; NID:g2367111; PIDN:AA
Mol. Gen. Genet. 229, 341-352, 1991
R;Dassa, J.; Faihi, H.; Marck, C.; Dion, M.; Kieffer-Bontemps, M.; Boquet, P.L.
A;Title: A new oxygen-regulated operon in Escherichia coli comprises the genes for a put
A;Reference number: S17958; MUID:92049231; PMID:1658595
A;Accession: S17960
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-17 <DA2>
A;Cross-references: UNIPARC:UPI00016FFB1; GB:S63811; NID:g238656; PIDN:AAB20286.1; PID:.

R.;Greiner, R.; Konietzny, U.; Jany, K.D.
Arch. Biochem. Biophys. 303, 107-113, 1993
A;Title: Purification and characterization of two phytases from Escherichia coli.
A;Reference number: S33278
A;Molecule type: protein
A;Accession: S33278; MUID:93256556; PMID:8387749
A;Residues: 23-31,'A',33-34 <GR2>
A;Cross-references: UNIPARC:UPI000179816
A;Comment: In addition to cAMP-mediated control, this enzyme is induced when bacterial o
naerobic conditions.
C;Genetic:
C;Gene: appA
A;Map position: 25 min
C;Keywords: monomer; periplasmic space; phosphohistidine; phosphoprotein; phosphoric mon
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-112/Product: acid phosphatase #status predicted <MAT>
F;38/Active site: Arg #status predicted
F;39/Active site: His (phosphohistidine intermediate) #status predicted

```
Query Match     98.1%;  Score 2182;  DB 2;  Length 432;
Best Local Similarity   96.6%;
Matches  424;  Conservative  0;  Mismatches  8;  Indels  0;  Gaps  0;

Qy    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60
         ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MKAILLIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60

Qy   61  TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE  120
         |||||||| ||||||||||||||| |||||||| |||||||||||||||||||  |||||
Db   61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDBRTRKTGE  120

Qy  121  APAAGLAPDCAITVHTQADTSSPDELFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH  180
         |||||||||||||||||||||||| |||||||||||||||||||||| |||||||||||
Db  121  APAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH   180

Qy  181  YQTAFRELERVLNFPQSNLCLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT  240
         ||||||||||||||||||||| ||||||||||||||||||||||| ||||||||||||||
Db  181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240

Qy  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA  300
         ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA  300

Qy  301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW    360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW   360

Qy  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420

Qy  421  IVNEARIPACSL  432
         ||||||||||||
Db  421  IVNEARIPACSL  432
```

RESULT 2

H90770
phosphoanhydride phosphorylase [imported] - Escherichia coli (strain O157:H7, substrain
phosphoanhydride phosphorylase
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: H90770
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kihara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: H90770
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-434 <HAY>
A;Cross-references: UNIPROT:Q8XC29; UNIPARC:UPI00000D0517; GB:BA0000007; PIDN:BAB34559.1;
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs1136

```
Query Match     95.0%;  Score 2144;  DB 2;  Length 434;
Best Local Similarity   96.5%;
Matches  419;  Conservative  0;  Mismatches  13;  Indels  2;  Gaps  1;

Qy    1  MKAILIPFLSLLIPLTPQSAFAQS--EPELKLESVVIVSRHGVRAPTKATQLMQDVTPDA   58
         ||||| ||||||||||||||||||  |||||||||||||||||||||||||||||||||
Db    1  MKAILLIPFLSLLIPLTPQSAFAQSEPEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDA  60

Qy   59  WPNWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLTKKGCPQPGQVAIIADVDERTRKT  118
         ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   61  WPNWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLTKKGCPQPGQVAIIADVDERTRKT  120

Qy  119  GEAPAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFT   178
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GEAPAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFT   180

Qy  179  GHYQTAFRELERVLNFPQSNLCLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASM  238
         || ||||||||||||||||||| ||||||||||||||||||||||| |||||||||||||
Db  181  GHRQTAFRELERVLNFPQSNLCLNREKQDESCSLTQALPSELKVSADNVSLTGAVSLASM  240

Qy  239  LTEIFLLQQAQGMPEPGWGRITDSHQWNTLLLSLHNAQFDLLQRTPEVARSRATPLLDLIK  298
         ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db  241  LTEIFLLQQAQGMPEPGWGRITDSHQWNTLLLSLHNAQFYLLQRTPEVARSRATPLLDLIM  300

Qy  299  TALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFE   358
         | |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  IALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFE   360

Qy  359  RWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGF  418
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  RWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGF  420

Qy  419  TQIVNEARIPACSL  432
         ||||||||||||||
Db  421  TQIVNEARIPACSL  434
```

RESULT 3

D85633
hypothetical protein appA [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: D85633
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D85633
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-444 <STO>
A;Cross-references: UNIPROT:Q8XC29; UNIPARC:UPI000165751; GB:AE005174; NID:g12514245; P
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: appA

```
Query Match     96.5%;  Score 2144;  DB 2;  Length 444;
Best Local Similarity   96.5%;
Matches  419;  Conservative  0;  Mismatches  13;  Indels  2;  Gaps  1;

Qy    1  MKAILIPFLSLLIPLTPQSAFAQS--EPELKLESVVIVSRHGVRAPTKATQLMQDVTPDA   58
         ||||| ||||||||||||||||||  |||||||||||||||||||||||||||||||||
Db   11  MKAILLIPFLSLLIPLTPQSAFAQSEPEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDA  70

Qy   59  WPFTWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKT  118
         ||  ||||||| ||||||||||||| |||||||| ||||||||||||||||||||||||
Db   71  WPNWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLTKKGCPQPGQVAIIADVDERTRKT  130

Qy  119  GEAPAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFT   178
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  178  GEAPAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFT   178
```

```
Qy  131 GEAFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFT 190
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  179 GHYQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASM 238
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  191 GHRQTAFRELERVLNFPQSNLCLNREKQDESCSLTQALPSELKVSADNVSLTGAVSLASM 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  239 LTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIK 298
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  251 LTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYILQRTPEVARSRATPLLDLIM 310
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  299 TALTHPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPE 358
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  311 IALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPE 370
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  359 RWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGF 418
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  371 RWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGF 430
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  419 TQIVNEARIPACSL 432
        ||||||||||||||
Db  431 TQIVNEARIPACSL 444
```

RESULT 4

AC0201
acid phosphatase (EC 3.1.3.2) [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AC0201
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AC0201
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-441 <KUR>
A;Cross-references: UNIPROT:Q82FP6; UNIPARC:UPI00000DCCF7; GB:AL590842; PIDN:CAC90470.1;
A;Gene: YPO1648
C;Genetics:
C;Keywords: phosphoric monoester hydrolase

```
Query Match   42.5%;  Score 959.5;  DB 2;  Length 441;
Best Local Similarity   46.1%;  Pred. No. 3.3e-69;
Matches 201;  Conservative  67;  Mismatches 153;  Indels  15;  Gaps  7;

Qy    5 LIPPLSLIPLTPQSAFAQSEPB-LKLESVVIVSRHGVRAPTKATQLMQDVTPDAWPTWP 63
          |:: |||||||||| : ||||||||||||| :||||
Db   13 LVLMLSGLAAITAPVA---ABPSGYTLERVVILSRHGVRSPIKQTQLMNDVTPDKWPQWP 69

Qy   64 VKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGEAFA 123
          |: |||:| |||: ::|||  :: |||| ::|||| |||  || :|:: || |
Db   70 VKAGYLTPRGAELVTLMGGFYGDYFRSLGLL-AAGCPAEGGVYAQADIDQRTRLTGQAFL 128

Qy  124 AGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGHYQT 183
          ||| | |:| || ||||||||| |||:|:|| || ::  |||| ||||||||| ||:|||
Db  129 DGVAPGGLTVHNQADLKKTDLPLPHPVEAGVCKLDAAQTDKAIEEQLGGPLDTVSQRYAK 188

Qy  184 APRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSAD--CVSLTGAVSLASMLTE 241
          ||| |||||||||||||||||||||||||||||||||||:|    ||||||||||||:|
Db  189 PFAQMGDVLNFAASPYCKSLQQGKTCDFAHFAANEVNVNKEGTKVTLSGPLALSSTLGE 248

Qy  242 IFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRAATPLLDLIKTAL 301
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db  249 IFLLQNAQAMPEVAWQRLKGAENWVSLLSLHNAQFNLMAKTPYIARHGTPLLLQQIDTAL 308

Qy  302 TPHPPQKQAYGVTLPTS----VLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVF 357
          ||||||||||||||||     |||||||||||||||||||||||||||||||||||||||
Db  309 T---LQLDAQGQKLPISAQNRVLFLGGHDTNIANIAGMLGADWQLPEQPDNTPPGGGLVF 365
```

```
Qy  358 ERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSL-NTPPGEVKLTLAGCEERNAQGMCSLA 416
          : |||| ::  |||
Db  366 ELWQNPDNHQRYVAVRKMFYQTMDQLRNAEKLIDLKNNPAGIISVAVAGCENNGGDDKLCELD 425

Qy  417 GFTQIVNEARIPACSL 432
          |  |:|
Db  426 TFQKVAKVIEPACHI 441
```

RESULT 5

F90773
periplasmic glucose-1-phosphatase [imported] - Escherichia coli (strain O157:H7, substra
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: F90773
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F90773
A;Molecule type: DNA
A;Residues: 1-413 <HAY>
A;Cross-references: UNIPROT:Q8XBZ6; UNIPARC:UPI00000D0533; GB:BA000007; PIDN:BAB34581.1;
A;Experimental source: strain O157:H7, substrain RIMD 0509952
A;Genetics:
A;Gene: ECs1158

```
Query Match   24.8%;  Score 560;  DB 2;  Length 413;
Best Local Similarity   33.3%;  Pred. No. 3.7e-37;
Matches 141;  Conservative  73;  Mismatches 174;  Indels  36;  Gaps  12;

Qy   13 IPLTPQSAFAQSEPB-LKLESVVIVSRHGVRAP-TKATQLMQDVTPDAWPTWPVKLGELT 70
          :| ||||| :  ||
Db   13 VVLLASNAQAQTVPEGYQLQQVLMMSRHNLRAPLANNGSVLEQSTPNKWPEWDVPGGQLT 72

Qy   71 PRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGEAFAAGLAPDC 130
          ||| | :|:|||: :| |   ||| |||| |||: |:| :|  || |
Db   73 TKGGVLEVYMGHYHMREWLAEOGMVVKSGECPPPDTVYAYANSLQRTVATAQFFITGAFPGC 132

Qy  131 AITVHTQADTSSPDLFNPLKTGVCQLDNANVTD----AILERAGGSIADFTGHYQTAFRE 187
          :| || |: | |
Db  133 DIPVHHQEKMGTMDPTPNPVITD-----DSAAAFSEQAVAAMEKLSKL-QLTDSYQL--- 183

Qy  188 LERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLTEIFLLQQ 247
          || |:|: | ||     |:| |||| ||  | |
Db  184 LEKIVNYKDSPAC-----KEKQQCSLVDGKNTFSAKYQQBPGVSGPLKVGNSLVDAFTLQY 239

Qy  248 AQGMP--EPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRAATPLLDLIKTALTHP 305
          | ||   :  |   ||| || ||| ||  ||| | |||||  | |
Db  240 YEGFPMDQVAWGEIKSDQQWKVLSKLRNGYQDSLFTSPEVARNVAKPLVSYIDKAL--- 295

Qy  306 PQKQAYGVTLPTS---VLFIAGHDTNLANLGGALELN-WTLPGQPDNTPPGGELVFERWR 361
                || :|  |||||  | || |:  |  ||  :|| |||| :|||||| :||| |
Db  296 ------VTDRTSAPKITVLVGHDSNIASLLTALDFKPYQLHDQNERTPIGGKIVFQRVR 348

Qy  362 RLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQI 421
          ::|:| |:| | | |  | ||  || | |   |  ||| ||||  |  |:| |   | :
Db  349 DSKANRDLMKIEYVVYQSAEQLRNADALTLQAPAQRVTLELSGC-PIDADGFCPMDKFDSV 407

Qy  422 VNEA 425
          :|||
Db  408 LNEA 411
```

RESULT 6

B85636
periplasmic glucose-1-phosphatase [imported] - Escherichia coli (strain O157:H7, substra
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: B85636
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
Nature 409, 529-533, 2001
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: B85636
A;Map position: 23 min
C;Genetics:
A;Gene: agp
C;Function:
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-413 <STO>
A;Cross-references: UNIPROT:Q8XBZ6; UNIPARC:UPI00000D0533; GB:AE005174; NID:g12514273;
C;Keywords: homodimer; periplasmic space; phosphoprotein; phosphoric mo
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: agp

Query Match   24.8%; Score 560; DB 2; Length 413;
Best Local Similarity 33.3%;   Pred. No. 3.7e-37;
Matches 141; Conservative 73; Mismatches 174; Indels 36; Gaps 12;

```
Qy   13 IPLTPQSAFAQSEPE-LKLESVVIVSRHGVRAP-TKATQLMQDVTPDAWPTWPVKLGELT  70
            : :| | |  : ||: |:||| :  |||||:||  ||  | |  |  ||  ||| :| :||
Db   13 VVLLASNAQAQTVPEGYQLQQVLMMSRHNLRAPLANNGSVLEQSTPNKWPEWDVPGGQLT  72

Qy   71 PRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGEAFAAGLAPDC 130
            |||| |  |:|:|:||  | :|::  | ||| |||||||:|||||:|||||:|:  ||  :
Db   73 TKGGVLEVYMGHYMREWLAEQGMVKSGECPPPDTVVAYANSLQRTVATAQFFITGAFPGC 132

Qy  131 AITVHTQADTSSPDPLFNPLKTGVCQLDNANVTD---AILERAGGSIADFTGHYQTAFRE 187
            :| ||:: |    ||:|||::| :  : | ::|:|    ::::  :: |::: ||:|::
Db  133 DIPVHHQEKMGTMDPTFNPVITD----DSAAFSEQAVAAMEKELSKL-QLTDSYQL---- 183

Qy  188 LERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLTEIFLLQQ 247
            ||::: : :|  ||  |:|:  ||: : |  ||   ||||| :|    |::  :: | |
Db  184 LEKIVNYKDSPAC----KEKQQCSLVDGKNTFSAKYQQEPGVSGPLKVGNSLVDAFTLQY 239

Qy  248 AQGMP--EPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTALTPHP 305
             :| |    ||||  | | ||  | : |  || | :::||||| | ||:   | : |
Db  240 YEGFPMDQVAWGEIKSDQQWKVLSKLKNGYQDSLFTSPEVARNVAKPLVSYIDKAL---- 295

Qy  306 PQKQAYGVTLPTS---VLFIAGHDTNLANLGGALELN-WTLPGQPDNTPPGGELVFERWR 361
                   | : | |   :|:: |||:|:| | :|  | : | :    ::| ::||:: ||
Db  296 -------VTDRTSAPKITVLVGHDSNIASLLTALDFKPYQLHDQNERTPIGGKIVFQRWR 348

Qy  362 RLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQI 421
            : : |  :::|  : |:::|:|::  ::| | |  ::  | |||   : |  | |||:: :
Db  349 DSKANRDLMKIEYVYQSAEQLRNADALTLQAPAQRVTLELSGC-PIDADGFCPMDKFDSV 407

Qy  422 VNEA 425
            :|||
Db  408 LNEA 411
```

A;Cross-references: UNIPROT:UPI000001256FB; GB:AE000202; GB:U000096; NID:g1787233; PIDN:AAC
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: agp
C;Function:
A;Description: essential for growth in a high-phosphate medium containing glucose-1-phosp
A;Note: optimal at low pH
C;Keywords: homodimer; periplasmic space; phosphoprotein; phosphoric mo
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-413/Product: glucose-1-phosphatase #status predicted <MAT>
F;40/Active site: His (phosphohistidine; phosphoric mo
F;311/Active site: His #status predicted

RESULT 7
JV0087
glucose-1-phosphatase (EC 3.1.3.10) precursor - Escherichia coli (strain K-12)
C;Species: Escherichia coli
A;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
A;Accession: JV0087; H64841
R;Pradel, E.; Marck, C.; Boquet, P.L.
J. Bacteriol. 172, 802-807, 1990
A;Title: Nucleotide sequence and transcriptional analysis of the Escherichia coli agp ge
A;Reference number: JV0087; MUID:90130318; PMID:2153660
A;Accession: JV0087
A;Molecule type: DNA
A;Residues: 1-413 <PRA>
A;Cross-references: UNIPROT:P19926; UNIPARC:UPI000001256FB; GB:M38807; NID:g145217; PIDN:
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A;Title: The complete genome sequence of Escherichia coli K-12.
Science 277, 1453-1462, 1997
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: A64720
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-413 <BLAT>

A;Cross-references: UNIPROT:UPI000001256FB; GB:AE000202; GB:U000096; NID:g1787233; PIDN:AA

RESULT 8
AG0632
glucose-1-phosphatase precursor (Glpase), secreted [imported] - Salmonella enterica subs
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
A;Note: this species has also been called Salmonella typhi
A;Accession: AG0632
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AG0632
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-413 <PAR>
A;Cross-references: UNIPARC:UPI000000059FE6; GB:AL513382; PIDN:CAD08242.1; PID:g16550289;
C;Genetics:
A;Gene: STY1153

Query Match   24.5%; Score 554; DB 2; Length 413;
Best Local Similarity 33.3%;   Pred. No. 1.1e-36;
Matches 141; Conservative 72; Mismatches 175; Indels 36; Gaps 12;

```
Qy   13 IPLTPQSAFAQSEPE-LKLESVVIVSRHGVRAP-TKATQLMQDVTPDAWPTWPVKLGELT  70
            : :| | |  : ||: |:||| :  |||||:||  ||  | |  |  ||  ||| :| :||
Db   13 IVLLASNAQAQTVPEGYQLQQVLMMSRHNLRAPLANNGSVLEQSTPNKWPEWDVPGGQLT  72

Qy   71 PRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGEAFAAGLAPDC 130
            |||| |  |:|:|:||  | :|::  | ||| |||||||:|||||:|||||:|:  ||  :
Db   73 TKGGVLEVYMGHYMREWLAEQGMVKSGECPPPDTVVAYANSLQRTVATAQFFITGAFPGC 132

Qy  131 AITVHTQADTSSPDPLFNPLKTGVCQLDNANVTD---AILERAGGSIADFTGHYQTAFRE 187
            :| ||:: |    ||:|||::| :  : | ::|:|    ::::  :: |::: ||:|::
Db  133 DIPVHHQEKMGTMDPTFNPVITD----DSAAFSEQAVAAMEKELSKL-QLTDSYQL---- 183

Qy  188 LERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLTEIFLLQQ 247
            ||::: : :|  ||  |:|:  ||: : |  ||   ||||| :|    |::  :: | |
Db  184 LEKIVNYKDSPAC----KEKQQCSLVDGKNTFSAKYQQEPGVSGPLKVGNSLVDAFTLQY 239

Qy  248 AQGMP--EPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTALTPHP 305
             :| |    ||||  | | ||  | : |  || | :::||||| | ||:   | : |
Db  240 YEGFPMDQVAWGEIKSDQQWKVLSKLKNGYQDSLFTSPEVARNVAKPLVSYIDKAL---- 295

Qy  306 PQKQAYGVTLPTS---VLFIAGHDTNLANLGGALELN-WTLPGQPDNTPPGGELVFERWR 361
                   | : | |   :|:: |||:|:| | :|  | : | :    ::| ::||:: ||
Db  296 -------VTDRTSAPKITVLVGHDSNIASLLTALDFKPYQLHDQNERTPIGGKIVFQRWH 348

Qy  362 RLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQI 421
            : : |  :::|  : |:::|:|::  ::| | |  ::  | |||   : |  | |||:: :
Db  349 DSKANRDLMKIEYVYQSAEQLRNADALTLQAPAQRVTLELSGC-PIDADGFCPMDKFDSV 407

Qy  422 VNEA 425
            :|||
Db  408 LNEA 411
```

RESULT 9
S25627

glucose-1-phosphatase (EC 3.1.3.10) precursor - Providencia rettgeri
C;Species: Providencia rettgeri
C;Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S25627
R;Riccio, M.L.; Chiesurin, A.; Lombardi, G.; Satta, G.
submitted to the EMBL Data Library, September 1992
A;Reference number: S25627
A;Molecule type: DNA
A;Residues: 1-417 <RIC>
A;Cross-references: UNIPROT:Q52309; UNIPARC:UPI000012S6FC; EMBL:X68201; NID:g45771; PIDN
A;Experimental source: strain PV7
C;Genetics:
A;Gene: agp
C;Keywords: periplasmic space; phosphohistidine; phosphoprotein; phosphoric monoester hy
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-417/Product: glucose-1-phosphatase #status predicted <MAT>
F;42/Active site: His (phosphohistidine intermediate) #status predicted
F;314/Active site: His #status predicted

```
Query Match             23.8%;  Score 538;  DB 2;  Length 417;
Best Local Similarity   30.4%;  Pred. No. 2.2e-35;
Matches 130;  Conservative   85;  Mismatches 175;  Indels  38;  Gaps  11;

Qy     10 SLLIPLTPQSAFQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWPTWPVK 69
          :||||||| ||||||:|||:||||||||||||||||||:|||||||||||||||||
Db     13 ALFAPIAPTMASTDNQADMVLDQVLVSRHGVRHNLRTPIVNTGILTEVTDKKWPDWDAKSGYL 72

Qy     70 TPRGGELIAYLGHYWRQ-----RLVADGLLPKCGCPQSGQO-VAIIADVDERTRKTGEAPA 123
          ::||||||||||:||||     |:||||||||||||||||| :||:|||||||||||||||
Db     73 TTQGGALEVYMGHYFREWIDQNRLLADEL----CPTSNEDIYLYTNSLQRTIATAQFFA 127

Qy    124 AGIAPDCAITVHTQADTSSPDFLFNPLKTGVCQLDNANVTDAILERAGGSIADFTG--HY 181
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Db    128 AGAPPCCKVNIHHQPEIGKMDPVPNPIIT------NGSPEFKQKALAANDDYLKGISL 179

Qy    182 QTAFRELERVLNFPQQSNLCLKREKQODESCSLTQALPSELKVSADCVSLTGAVSLASMLTE 241
          :| |||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db    180 KAGYBELDTVLNIKDSQKC---KTDKLCNLDSQKNSFIIBADKEPGVSGPLKIANSAVD 235
```

```
Qy    242 IFLLQQAQGMP--BPGWGRITDSHQWNTLLSLHNAQPDLLQRTPEVARSRATPLLDLIKT 299
          ||||||||||  ||||||||||||||||||||||||||||||||||||||||||||:||
Db    236 AIDLQYYEGEPADQVAWGLVDTPEKWKGLNTLRNAYQEFLPTPKIIAKNVAHPILNYIDK 295

Qy    300 ALFEHPPQOKQAYGVTLPISVLFIAGHDTNLANLGGALELN-WTLPGQPDNTPPGGELVPE 358
          |:|||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db    296 GFVSDVKGETA------KFIFLVGHDSNIASLMSAMDPKYQLAQQYEHTPIGGKLVPQ 348

Qy    359 RWRRLSDNSQWIQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCE-ERNAQGMCSLAG 417
          ||||||||||||||||||||||||||||||||||||||||||||||| |:|||||||||
Db    349 RWTDKQTKXDFMKVEYVYQTADQLRDNAYLSLETPPKHVTLELKDCPVDKN--GYCSWED 406

Qy    418 FTQIVNBA 425
          ||:|||:|
Db    407 FQKWMAKA 414
```

RESULT 10
E87316

periplasmic phosphoanhydride phosphohydrolase [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: E87316
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Brmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: E87316
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-414 <STO>
A;Cross-references: UNIPROT:Q9AAQ4; UNIPARC:UPI00000C70A0; GB:AE005673; NID:g13421731; P
C;Genetics:
A;Gene: CC0542

```
Query Match             23.7%;  Score 536;  DB 2;  Length 414;
Best Local Similarity   33.6%;  Pred. No. 3.1e-35;
Matches 145;  Conservative   62;  Mismatches 186;  Indels  38;  Gaps  11;

Qy      9 LSL---LIPLTPQSAFQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWPTWPVK 65
          ||     ::||||||:|||||||||||||||||||||||||||||||:|||||||||||
Db      3 LSLRIAVVTLALASAGAASAB--TLEKVVILSRHGVRSAMSSPERLEBASARPNPRFEVP 60

Qy     66 LGEELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQOVAIIADVDERTRKTGEAPAAG 125
          |||:|||||||||||||||||||:|||||||   |||||||||||||||||||||||||
Db     61 AGHLTARGETLVARMGDYYRRHYAAQGLLKPGDC---ASVYAWANVQRTIATAKAYRET 117

Qy    126 LAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGHYQTAF 185
          |||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db    118 LAPGCPVTNTVGB-GNIDPMFEPVKAGIVKADHALARAAVAAGRVGGDLTAWSASHNQEA 176

Qy    186 RELERVLNFPQQSNLCLKREKQODESCSLTQALPSELKV-------SADCVSLTGAVSLA 236
          |:|||||||||||||||||||||||||||||||||||||       :|||||||||||
Db    177 EQLDALLMQCDKGPC------------PPAPGKRRVFDAKPGFVDGBELAGLSGPEAPA 223

Qy    237 SMLTBIFLLQQAQGMPBPGWG-RITDSHQWNTLLSLHNAQPDLLQRTPEVARSRATPLLD 295
          ||||:||||||||||  ||||  |||||||||||||||||||||||||||||||||:||
Db    224 SGVTESLLMAWADGRDPAGLGWKSLDEEALTRSFFLHQAEFDLRLRTPYVARTLAGHLAD 283

Qy    296 LIKTALPIPHPPQOKQAYGVTLPISVLFIAGHDTNLANLGGALELNWTLPG-QPDNTPPGGE 354
          ||:|||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    284 RLAATLR---DGAAAIG-PVDARLVIIAGHDGTLASLGGLIREWTLPGYQPNQIQPGGA 339

Qy    355 LVPERWRRLSDNSQWIQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCS 414
          ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||:
Db    340 LVFERWRR-DDGVRVRVRFTGQSLSQLRNMTALDAKTPPLSAPVFVQGCGTIATPAFDCR 398

Qy    415 LAGFTQIVNBA 425
          |:||:|||:|
Db    415 LAGFTQIVNBA 425
```

Db   399 LEDPETVVRGA 409

RESULT 11

```
T16058
hypothetical protein F13D11.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
A;Accession: T16058
R;Fulton, L.
A;Reference number: S69020
A;Molecule type: DNA
A;Residues: 1-416 <FUL>
A;Cross-references: UNIPARC:UPI000007AD42; EMBL:U40939; NID:g1072175; PI
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Accession: T16058
A;Residues: 23/3; 59/3; 88/3; 138/3; 202/3; 244/1; 270/3; 312/3; 349/3; 377/1
A;Cross-references: UNIPROT:Q19390; UNIPARC:UPI000007AD42; EMBL:U40939; NID:g1072175; PI
C;Genetics:
A;Gene: CESP:F13D11.1
A;Introns: 23/3; 59/3; 88/3; 138/3; 202/3; 244/1; 270/3; 312/3; 349/3; 377/1
C;Superfamily: mammalian acid phosphatase
```

```
Query Match        6.1%;  Score 137; DB 2;  Length 416;
Best Local Similarity  21.7%;  Pred. No. 0.0035;
Matches  99;  Conservative 61;  Mismatches 166;  Indels 130;  Gaps 21;

Qy    1 MKAILIPFLSLLIPLTPQSAPAQSPELKLESVVIVSRHGVRAPTKATQLMQDVTPD---  57
          :|::|:|  ||:||  |::|:    ||||:
Db    1 MRLVLLLP--FLFPV---AFG----KLKFVQIWFRHGERTPGHYLYFPGDDLNNVDY   48

Qy   58 ---AWPTWPVKLGELTPRGGELIAYLGHYWRQRL-----VADGLLPKCCPQSGQVAI 107
Db   49 QQIAWP-------GELTKRGILEEFQLG------ORLRKIYGEHFGDTYQPR-----DFHV  91

Qy  108 IADVDERTRKTGEAPAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAIL 167
Db   92 YTGKDNRTSASAQAMPAGFLPP-----NEDQTWNVELKWQPV-----AQLTDESI 136

Qy  168 ERAGGSIADFTGHYQTAFRE-----LERVILNFPQSNLCLKREKQDESCSLTQALPSEL 220
Db  137 DWVSLGAIDNCPVYGEAQRKSSEYAEVMDQMEKYDABLLQLVRNHADE-----  184

Qy  221 KVSADCVSLTGAVSLASMLTEI-----FLLQQAQGMEPEPGWGRITDSHQWNTLLSLHNAQPD 277
Db  185 PIVEAVKNHVIDSLKVRYILQDDR-LPYPEWARGYENRILNMSFLIHDA---  233

Qy  278 LLQRPTEVARSRATPLLD-----LIKTALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLG 333
Db  234 ---VVKUQNDSVGDYHNELVMSYFETHLQKSTKGV-----PISGHDTNLVTIW 279

Qy  334 GALELNWTLPCQPDNTPP-GGELVERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNT 392
Db  280 ESLRLD-----GHPEDIPNYGAHIAIE-----MHEPVGQLSIKPFLSMGF 319

Qy  393 PPGEVKLTLLAGCEERNAQGMCSLAGFTQIVNEARIP 428
Db  320 NQTRVELHPHFC-SRSQNNDCTWDEFQRLVKKSRKP 354
```

```
A;Cross-references: UNIPROT:P11117; UNIPARC:UPI000013IFE1; EMBL:X15525; NID:g34239; PIDN
R;Geier, C.; von Figura, K.; Pohlmann, R.
Eur. J. Biochem. 183, 611-616, 1989
A;Title: Structure of the human lysosomal acid phosphatase gene.
A;Reference number: S05525; MUID:89377828; PMID:2776754
A;Accession: S05525
A;Molecule type: DNA
A;Residues: 1-29 <GEI>
A;Cross-references: UNIPARC:UPI0000I727CD
R;Pohlmann, R.; Krentler, C.; Schmidt, B.; Schroeder, W.; Lorkowski, G.; Culley, J.; Mers
EMBO J. 7, 2343-2350, 1988
A;Title: Human lysosomal acid phosphatase: cloning, expression and chromosomal assignment
A;Reference number: S01155; MUID:89052645; PMID:3191910
A;Accession: S01155
A;Molecule type: mRNA
A;Residues: 1-423 <POH>
A;Cross-references: UNIPARC:UPI000013IFE1; EMBL:X12548; NID:g34262; PIDN:CAA31064.1; PID
A;Note: part of this sequence, including the amino end of the mature protein, was confir
C;Genetics:
A;Gene: GDB:ACP2
A;Cross-references: GDB:118963; OMIM:171650
A;Map position: 11p11.2-11p11.11
A;Introns: 38/3; 70/3; 99/3; 150/3; 183/3; 213/3; 258/3; 285/3; 321/2; 380/1
C;Superfamily: mammalian acid phosphatase
C;Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydrolas
C;Function: catalyzes the hydrolysis of a wide range of phosphate esters
A;Description: signal sequence
F;1-30/Domain: signal sequence #status predicted <SIG>
F;31-423/Product: acid phosphatase ACP2 #status experimental <MAT>
F;41/Active site: Arg #status predicted
F;42/Active site: His (phosphohistidine intermediate) #status predicted
F;92,133,167,177,191,267,322,331/Binding site: carbohydrate (Asn) (covalent) #status pre
F;159-370,212-310,345-349/Disulfide bonds: #status predicted
```

```
Query Match        5.8%;  Score 132; DB 1;  Length 423;
Best Local Similarity  23.3%;  Pred. No. 0.0091;
Matches  97;  Conservative 43;  Mismatches 173;  Indels 104;  Gaps 20;

Qy   30 LESVVIVSRHGVRAPTKATQLMQDVTPDAWPTWPVKLGELTPRGGELIAYLGHYWRQRLV  89
Db   33 LRFVTLLYRHGDRSPVKT----YPKDPYQEEEWPQGFGQLTKEGMLQHWELGQALRQRY--  87

Qy   90 ADGLLPKCCGCPQSGQVAIIADVD-ERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPL--  146
Db   88 -HGFL---NTISYHRQEVYVRSTDFDRTLMSAEANLAGLFP-----PNGMQR 129

Qy  147 FNP-LKTGVCQLDNANVTDAILERAGGSIADFTGHYQTAFRELERVILNFPQSNLCLKREK 205
Db  130 FNPNISWQPIPVHTVPITE-----DRLLKFPLGPCPRYEQL 165

Qy  206 QDESCSLTQALPSELKVSADCVSLTGAVSLASMLTEIFL-----LQQAQGMEPEP 254
Db  166 QNE----TRQTPEYQNESSRNAQFLDMVANETGLTLDLTLLETVWNVVDTLFCEQTHGLRLP 221

Qy  255 GWGRITDSHQWNTLLSLHNAQPDLL----QRTPEVARSRATPLLDLIKTALTPHPPQKQAY 311
Db  222 PW----ASPQTMQRLSRLKDFSFRFLFGIYQQAEKARLQGGVLLAQIRKNLTLMATTSQ-- 276

Qy  312 GVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQ 371
Db  277 ---LP-KLLVYSAHDTTLVALQMALDVY----NGEQAPYASCHIFELYQEDSGN--FS 324

Qy  372 VSLVFQTLQQMRDKTPLSLNTPPGEVKLTLLAGCEERNAQGMCSLAGFTQIVNEARIP 428
Db  325 VEMYF--RNESDKAFWPLSLP----------GCPHR-------CPLQDFLRL-TEPVVP 363
```

RESULT 12

```
S06167
acid phosphatase (EC 3.1.3.2) ACP2 precursor [validated] - human
N;Alternate names: acid phosphatase, lysosomal
C;Species: Homo sapiens (man)
C;Date: 04-Dec-1992 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004
A;Accession: S06167; S05525; S01155
R;von Figura, K.
submitted to the EMBL Data Library, June 1989
A;Reference number: S06167
A;Accession: S06167
A;Molecule type: DNA
A;Residues: 1-423 <VON>
```

RESULT 13

```
A33395
acid phosphatase (EC 3.1.3.2) precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 29-Jan-1990 #sequence_revision 29-Jan-1990 #text_change 09-Jul-2004
C;Accession: A33395
```

R;Himeno, M.; Fujita, H.; Noguchi, Y.; Kono, A.; Kato, K.
Biochem. Biophys. Res. Commun. 162, 1044-1053, 1989
A;Title: Isolation and sequencing of a cDNA clone encoding acid phosphatase in rat liver
A;Reference number: A33395
A;Accession: A33395
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-423 <HIM>
A;Cross-references: UNIPROT:P20611; UNIPARC:UPI0000131FE2; GB:M27893; NID:g202933; PIDN:
C;Superfamily: mammalian acid phosphatase
C;Keywords: phosphohistidine; phosphoprotein; phosphoric monoester hydrolase
F;41/Active site: Arg #status predicted
F;1-33/Domain: signal sequence #status predicted
F;42/Active site: His (phosphohistidine intermediate) #status predicted

Query Match   5.2%;   Score 116.5;   DB 2;   Length 423;
Best Local Similarity   22.8%;   Pred. No. 0.16;
Matches 102;   Conservative 46;   Mismatches 185;   Indels 115;   Gaps 22;

Qy   2 KAILIPFL----SLLILPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAW 59
            ::| :| |||| ::|| : |||
Db   10 QAALLQFLLGMCLMVMPIQA-----RSLRPVTLLYRHGDRSPVKA-----YPKDPYQB 58

Qy   60 PTWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVD--ERTRKT 118
            : ||| :|| || | ||| ::: : |:| |||
Db   59 EKWPQGFGQLTKEGMLQHWELGQALRQRY--HGFL--NASYHROEVYVRSTDFDRTLMS 113

Qy   119 GEAPAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADPT 178
            :| ||| :   : || :||: :| ||:| : : : : ||
Db   114 AEANLAGLPFP-----TEVQHNPNISWQPI-----PVHTVPITB--------- 148

Qy   179 GHYQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLIGAVSLASM 238
            :| : : || ||| :| ::| : | : |:| |||
Db   149 ------------DRLLKFPLGPCPRYBQLONE-----TRCTPEYQNMSIQNAQPLDMVANETG 194

Qy   239 LTEIFL-----------LQQAQGMPEPGWGRITDSHQWNTLLSLHNAQP-------DLLQ 280
            |: || : ||
Db   195 LMNILLETIWNYYDTLFCEQTHGLLLPPW-----ASPQTVQALSQLKDFSFLFLFGIHDQVQ 251

Qy   281 RTPEVARSRATPLLDLIKTALTPHPPQKQAYGVTLPTSVLPIAGHDTNLANLGGALELNW 340
            :: || | | :||| || : | | | :
Db   252 K-----ARLQGGVILAQILKNLTLMATTSQF----------PKLLVSAHDTLVALQMALNVVY- 300

Qy   301 ----NGKQAPYASCHIFELYQ--BDNGNF-SVEMYFRNDSK----KAPW------PLT 341
            :| :| : ||: : : | |: ||
Db   341 TLPGQPDNTPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLT 400

Qy   342 LPGCPHR-------CPLQDFLRL-TEPVIP 363
            |||| | || |:|| |:|::|
Db   401 LAGCEERNAQGMCSLAGFTQIVNEARIP 428

RESULT 14
S64682
acid phosphatase (EC 3.1.3.2) Acph-1 precursor, lysosomal - fruit fly (Drosophila melano
C;Species: Drosophila melanogaster
C;Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 09-Jul-2004
C;Accession: S64682; S64681
R;Chung, H.J.; Shaffer, C.; MacIntyre, R.
Mol. Gen. Genet. 250, 635-646, 1996
A;Title: Molecular characterization of the lysosomal acid phosphatase from Drosophila me
A;Reference number: S64681; MUID:96194627; PMID:8676866
A;Accession: S64682
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-438 <CHU>
A;Cross-references: UNIPROT:Q9VAD0; UNIPARC:UPI000017595D
A;Accession: S64681
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-31,'Y',33-438 <CHW>
A;Cross-references: UNIPARC:UPI000017595B
A;Genetics:
A;Gene: FlyBase:Acph-1

Query Match   5.0%;   Score 112.5;   DB 2;   Length 438;
Best Local Similarity   20.3%;   Pred. No. 0.35;
Matches   82;   Conservative 59;   Mismatches 145;   Indels 117;   Gaps 19;

A;Cross-references: FlyBase:FBgn0000032
C;Superfamily: mammalian acid phosphatase
C;Keywords: phosphohistidine; phosphoric monoester hydrolase
C;Superfamily: mammalian acid phosphatase

Qy   29 KLESVVIVSRHGVRAPTKATQLMQDVTPDAWPT--------WPVKLGELTPRGGELIAYL 80
            :|: :||||:||||||||||||||||||| | ||||||||| ||||||
Db   47 QLKFVHVIYRHGDRTPV----------------DPYPTDPWGDRKFWPTGWGDLTNLGKQEHYDL 95

Qy   81 GHYWRQRLVADGLLPKCGCPQSGQVAIIADVD--ERTRKTNGEAFAAGL---APDC----130
            ||:|||||||:||||||||||||||||||||| ||||| :|||||| ||
Db   96 GKWLRNRY--SNLLPPI--YSNENIYVQSTDVDRTLMSAQSNLAGLYEPQGEDIWNTDI 150

Qy   131 ---AITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADPTGHYQTAFRE 187
              |:||||||||||||||||||||||||||||||||||:||||||| :|
Db   151 NWQPIPIHTSPEREDP----------------ILAAKARCPAYDY-----------E 180

Qy   188 LERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLIGAVSLASMLTEIFLLQQ 247
            |||||||||||||||||||||||||||||||||||||||||||||||||:|| |||:|
Db   181 LASLESSPEFK-ALTEKHRNLFAYLSEKGGRPVKTFID----------AQYLNNTLFIEN 229

Qy   248 AQGMPEPGWGRITDSHQWNTLLSLHNAQPDLLLQRTPEVARSRATPLL-DLIKTALTPHPP 306
            ||||||||||||||||||||||||||||||: |||||| ||||||| | :|||| ||
Db   230 LYNMTLPKMTKKVYGREELTYVS--NFAPAISSYTRKLARLKAGPLLKDIFQRF----281

Qy   307 QKQAYGVTLPTSVLPI-AGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERWRRLSD 365
            || | ||| ||| :| | ||| || |:|:| |||||||||||||| ||| |: :
Db   282 KEKSSGSLKPDRSMWVYSAHDTTVASVLNALKL-FEL------HSPP------321

Qy   366 NSQWIQVSLVFQTLQQMRDKTPL-----SLNTPGEVKLTLAGC 404
            |||:|||||||||||||:|||| ||| |||:||| ||
Db   322 -----YTACIMMELRVDETNTPLVSIFYKQNTTABPLPLDIPGC 359

RESULT 15
S14742
acid phosphatase (EC 3.1.3.2) precursor - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S14742
R;Geier, C.; von Figura, K.; Pohlmann, R.
Biol. Chem. Hoppe-Seyler 372, 301-304, 1991
A;Title: Molecular cloning of the mouse lysosomal acid phosphatase.
A;Reference number: S14742; MUID:91282986; PMID:2059937
A;Accession: S14742
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-421 <GEI>
A;Cross-references: UNIPROT:P24638; UNIPARC:UPI000002710J; EMBL:X57199; NID:g52870; PIDN
C;Superfamily: mammalian acid phosphatase
C;Keywords: lysosome; phosphoric monoester hydrolase

Query Match   4.9%;   Score 111.5;   DB 2;   Length 421;
Best Local Similarity   21.4%;   Pred. No. 0.4;
Matches   98;   Conservative 53;   Mismatches 171;   Indels 137;   Gaps 23;

Qy   2 KAILIPF-----LSLLILPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTP 56
            ::| :| |:::|| : : ||
Db   8 QAALLQFLLGMCLTVMPIQARS-----------------LRFVTLLYRHGDRSPVKT----YPKDP 53

Qy   57 DAWPTWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVD--ERT 115
            || |:||||||| ||||||||||||||||||||:||||||||||||||||||||| ||
Db   54 YQEEKWPQGFGQLTKEGMLQHWELGQALRQRY--HGFL---NTSYHROEVYVRSTDFDRT 108

Qy   116 RKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIA 175
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :
Db   109 LMSAEANLAGLPPP-----------------NSVQHFSPNISWQPI-----PVHTVPITE----146

```
Qy   176  DFTGHYQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSL  235
                                  :|:||       ||:
Db   147  ------------------------DRLLKFPLGPCRYEQLQNE-----TRQTPEYQNRSIQNAQFLNMVAN  189

Qy   236  ASMLTEIFL-------------------LQQAQGMPEGWG------RITDSHQWN--TLLSLHNAQP  276
             :|||:|:                 ::             ||           :|
Db   190  ETGLTNVTLETIWNVYDTLFCEQTHGLLLPPWASPQTVQRLSQLKDFSFLFGIHE---  246

Qy   277  DLLQRTPEVARSR--------ATPLLDLIKTALTPHPPQKQAYGVTLPTSVLPIAGHDTNL  329
             :|  ||:|          |   ||:||                    ||||  ||||:
Db   247  ------QVQKARLQGGVLLAQILKNLTLMATTSQFPK---------LLIVYSAHDTTL  288

Qy   330  ANLGGALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLS  389
             |||:|||:|         |:            ::      |:|:|
Db   289  VALQMALNVY-----NGKQAPYASCHIFELYQ--EDNGNF-SVEMYFRNDSK---KAPWP  337

Qy   390  LNTPPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIP  428
             ||||          |:  ||     |||    ||
Db   338  LILP---------GCPHR-----CPLQDFLRL-TEPVIP  361
```

Search completed: June 13, 2006, 10:29:30
Job time : 26.2661 secs

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

Run on: June 14, 2006, 10:58:44 ; Search time 5160.22 Seconds
                                   (without alignments)
                                   8030.264 Million cell updates/sec

Title: US-10-601-319-10
Perfect score: 2258
Sequence: 1 MKAILIPFLSLLLIPLTPQSA..........CSLAGFTQIVNEARIPACSL 432

Scoring table: BLOSUM62
                Xgapop 10.0 , Xgapext 0.5
                Ygapop 10.0 , Ygapext 0.5
                Fgapop   6.0 , Fgapext 7.0
                Delop   6.0 , Delext   7.0

Searched: 6366136 seqs, 3197310525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10601319/runat_12062006_180050_27480/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss06h
-USER=US10601319 @CGN_1.1.3991 @runat_12062006_180050 -DEV TIMEOUT=120
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
    1: gb_env:*
    2: gb_pat:*
    3: gb_ph:*
    4: gb_pl:*
    5: gb_pri:*
    6: gb_ro:*
    7: gb_sts:*
    8: gb_sy:*
    9: gb_un:*
   10: gb_vi:*
   11: gb_ov:*
   12: gb_htg:*
   13: gb_in:*
   14: gb_om:*
   15: gb_ba:*

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2182	96.6	1901	2	AR636190	AR636190 Sequence
2	2182	96.6	1901	2	AR636191	AR636191 Sequence
3	2182	96.6	1901	2	AX356572	AX356572 Sequence
4	2182	96.6	1901	15	AX356574	AX356574 Sequence
5	2182	96.6	1901	15	M58708	M58708 Escherichia
6	2182	96.6	3470	2	AX042376	AX042376 Escherichia
7	2182	96.6	4060	2	AX042374	AX042374 Sequence
8	2182	96.6	6116	2	AX042375	AX042375 Sequence
9	2181	96.6	6708	2	AX042373	AX042373 Sequence
10	2182	96.6	17732	2	AX042378	AX042378 Sequence
11	2182	96.6	20623	2	AX042372	AX042372 Sequence
12	2182	96.6	110000	15	AP009048_10	Continuation (11 o
13	2182	96.6	110000	15	CP000038_10	Continuation (11 o
14	2182	96.6	110000	15	AR636188	AR636188 Sequence
15	2182	96.6	1315	2	AX356188	AX356188 Sequence
16	2181	96.5	1901	15	AF537219	AF537219 Escherich
17	2178	96.5	1901	15	AX356570	AX356570 Escherichia
18	2176	96.4	1901	15	AF537219	AF537219 Escherichia
19	2176	96.4	1901	15	AX356570	AX356570 Sequence
20	2176	96.4	1901	15	L03370	ECOACPHC
21	2176	96.4	1901	15	L03372	ECOACPHA
22	2175	96.3	1901	15	L03370	ECOACPHD
23	2175	96.3	1901	2	L03374	ECOACPHE
24	2175	96.3	1901	2	AR636189	AR636189 Sequence
25	2172	96.2	1901	15	AX356571	AX356571 Sequence
26	2171	96.1	1323	2	L03371	ECOACPHB
27	2171	96.1	1323	2	L03375	ECOACPHF
28	2171	96.1	1323	2	AR108133	AR108133 Sequence
29	2171	96.1	1323	2	AR127818	AR127818 Sequence
30	2171	96.1	1323	2	AR130956	AR130956 Sequence
31	2171	96.1	1323	2	AR492966	AR492966 Sequence
32	2171	96.1	1323	2	AR636185	AR636185 Sequence
33	2171	96.1	1323	2	AX052713	AX052713 Sequence
34	2164	95.8	1486	2	AX338966	AX338966 Sequence
35	2154	95.4	1486	15	AX356566	AX356566 Sequence
36	2152	95.2	110000	15	AR630587	AR630587 Sequence
37	2149	95.2	5421	2	CP000036_22	Continuation (23 o
38	2147	95.1	1489	2	AX042377	AX042377 Sequence
39	2147	95.1	1489	15	AX042377	Continuation (11 o
40	2145	95.0	110000	15	AR277753	AR277753 Sequence
41	2144	95.0	110000	2	AY496073	AY496073 Escherich
42	2144	95.0	110000	2	AE005674_10	Continuation (11 o
43	2144	95.0	110000	15	AE014073_10	Continuation (11 o
44	2144	95.0	1486	15	AR630588	AR630588 Sequence
45	2143	94.9	110000	15	AE005174_12	Continuation (13 o

ALIGNMENTS

RESULT 1
AR636190

LOCUS   AR636190   1901 bp   DNA   linear   PAT 20-APR-2005
DEFINITION Sequence 7 from patent US 6855365.
ACCESSION   AR636190
VERSION   AR636190.1 GI:62768873
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1901)
  AUTHORS   Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B.,
            O'Donoghue,E. and Mathur,E.J.
  TITLE   Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: US 6855365-A 7 15-FEB-2005;
            Diversa Corporation; San Diego, CA
FEATURES   Location/Qualifiers
     source   1..1901
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:   3.74e-196
Score: 2182.00
Percent Similarity: 98.1%
Best Local Similarity: 98.1%

Length: 1901
Matches: 424
Conservative: 0
Mismatches: 8

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

```
Query Match:    96.6%    Indels:  0
DB:             2        Gaps:    0

US-10-601-319-10 (1-432) x AR636190 (1-1901)

Qy    1  MetLysAlaIleIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla  20
Db  188  ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCCGTTAACCCCGCAAATCGCA  247

Qy   21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly       40
Db  248  TTCGCTCAGAGTGAGCCCGAGCTGAAGCTGGAAAGTGTGGTTGTCAGTCGTCATGGT       307

Qy   41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspProAlaTrpPro  60
Db  308  GTGCGTGCTCCAACCAAAGCCACGCAACTGATGCAGGATGTCACCCCAGACCCAGCA       367

Qy   61  ThrTrpProValLysGlyGlyLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu      80
Db  368  ACCTGGCCGGTAAAACGGGGGTTGGCTGACACCCGCGGNGGTGGTGAGCTTATCGCCTATCTC  427

Qy   81  GlyHisTyrTrpArgGlnArgLeuValAlaAlaGlyLeuLeuProLysCysGlyCysPro   100
Db  428  GGACATTACCAACGCCAGCCGTCTGGTAGCCGCAGGACTGCTGCCGAAAAAGGGCTGCCCG  487

Qy  101  GlnSerGlyValAlaIleAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu  120
Db  488  CAGTCTGGTGTCGCGATTGCGATTATTGCTGACGTCGACGAGCGTACCCGTAAAACAGGCGAA  547

Qy  121  AlaPheAlaLeuAlaGlyLeuAlaAlaProAspCysAlaIleThrValHisThrGlnAlaThr  140
Db  548  GCCTTCGCCCTGGCCGGTCTGGCGGCACCTGACTGTGCGATTACCGTACATACCCAGGCGACC  607

Qy  141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAspAsnAla  160
Db  608  TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG  667

Qy  161  AsnValThrAspAlaIleLeuGluLysAlaGlyGlySerIleAlaAlaPheThrGlyHis     180
Db  668  AACGTGACTGACGCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGCTTTTACCGGCAT    727

Qy  181  TyrGlnThrAlaPheArgGluLeuGluGluGlnThrArgValLeuAsnProHeProGlnSerAsnLeuCys  200
Db  728  CGGCAAACGGCGTTTCGCGAACTGGAAGAGCTGCTTAATTTCCGCAATCAAAACTTGTGC  787

Qy  201  LeuLysArgGluLysLysGlnAspGluSerCysSerLeuThrGlnAlaIleLeuProSerGluLeu  220
Db  788  CTTAAACGTGAGAAAAACAAGGACGAAAGCTGTTCTCTCACCCAGGCAATTACCATCGGAACTC  847

Qy  221  LysValSerAlaAsnCysValSerLeuThrGlyAlaValSerIleLeuAlaSerMetLeuThr  240
Db  848  AAGGTGAGCGCCGACAATGTCATCAACCGTGCGGTAAGCCTCCCATCAATGCTGACG  907

Qy  241  GluIlePheLeuLeuGlnIleAlaGlyIleGlyMetProGlyLeuTrpGlyArgIleThr  260
Db  908  GAGATATTTCTCCTGCAACAACAGGCGAATCGGGGAGCGGGTTGGGGAAGGAATCACC  967

Qy  261  AspSerHisGlnTrpLeuAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln  280
Db  968  GATTCACACCAGTGGAACACTTGCTAAGTTTGCATAACGCCAATTTTATTGCTACAA  1027

Qy  281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla  300
Db  1028 CGCACGCCAGAGGTTGCGCGCAGCCGCGCCACCCCGTTATTAGATTTGATCAAAGACAGCG  1087

Qy  301  LeuThrProHisProProGlnIleLysSerGlnAlaTyrGlyValThrLeuProThrSerValLeu  320
Db  1088 TTGACGCCCATCCACCGCAAATAAAACAGGGTTNGGGGTGTGACATTACCAACTTCAGTGCTG  1147

Qy  321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340
Db  1148 TTTATCGCCGGACACGATACTAATCTCGGCGGCGGGCCACTGGAGCTCAACTGG  1207
```

```
US-10-601-319-10 (1-432) x AR636191 (1-1901)

Alignment Scores:
Pred. No.:              3.74e-196    Length:        1901
Score:                  2182.00     Matches:       424
Percent Similarity:     98.1%       Conservative:  0
Best Local Similarity:  98.6%       Mismatches:    8
Query Match:            96.6%       Indels:        0
DB:                     2           Gaps:          0

Qy  341  ThrLeuProGlyGlnProAspAsnThrProProGlyValGlyGluLeuValPheGlnAlaTrp  360
Db  1208 ACGCTTCCCGGTTCAGCCCGATAACACGCCGCCAGGTGGTGGTGAACTGGTTTTCCAGGCCTGG  1267

Qy  361  ArgArgLeuSerAspAsnSerGlnIleTrpIleGluGlnValSerLeuGlyValPheGlnArgThrLeuGln  380
Db  1268 CGTCGGCTAAGCGATAACAGCCAGATTTGGATTGAGCAGGTTTCGCTGGGTGTTCCGCAGACTTTACAG  1327

Qy  381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValArgLeuThr  400
Db  1328 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCCGGAGAGGTGAGACTGACC  1387

Qy  401  LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerSerLeuAlaGlyPheThrGln  420
Db  1388 CTGGCAGGATGTGAAGAGCGCAATGCCCAGGGCATGTGTTCGTCGTTTGCCGGTTTTACGCAA  1447

Qy  421  IleValAsnGluValAlaArgIleProAlaCysSerLeu  432
Db  1448 ATCGTGAATGAAGCACCGATACCGGCGTGCAGTTTG  1483
```

```
RESULT 3

LOCUS       AX356572                1901 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 7 from Patent WO0019033.
ACCESSION   AX356572
VERSION     AX356572.1  GI:18621059
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B. and
            O'Donoghue,E.
  TITLE     Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: WO 0190333-A 7 29-NOV-2001;
            DIVERSA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..1901
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

US-10-601-319-10 (1-432) x AX356572 (1-1901)

Alignment Scores:
Pred. No.:
Score:                    2182.00        Matches:           424
Percent Similarity:       98.1%          Conservative:        0
Best Local Similarity:    98.1%          Mismatches:          8
Query Match:              96.6%          Indels:              0
DB:                       2              Gaps:                0
```

```
Alignment Scores:
Pred. No.:              3.74e-196
Score:                  2182.00
Percent Similarity:     98.1%
Best Local Similarity:  98.1%
Query Match:            96.6%
```

```
RESULT 4
AX356574

LOCUS       AX356574               1901 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 9 from Patent WO0190333.
ACCESSION   AX356574
VERSION     AX356574.1  GI:18621060
KEYWORDS
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B. and
            O'Donoghue,E.
  TITLE     Recombinant bacterial phytases and uses thereof
  JOURNAL   Patent: WO 0190333-A 9 29-NOV-2001;
            DIVERSA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..1901
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN
```

```
DB: US-10-601-319-10          2          Gaps:          0

US-10-601-319-10 (1-432) x AX356574 (1-1901)

Qy    1  MetLysAlaIleIleLeuIlePheLeuSerLeuLeuIlePheProGlnSerAla   20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  188  ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTTATTTTTCCTCAAAGTGCA  247

Qy   21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly   40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  248  TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT  307

Qy   41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  308  GTGCGTGCGCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCCTGGCCA  367

Qy   61  ThrTrpProValLysGlyGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu   80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  368  ACCTGGCCGGTAAAAGGTGGTGAGCTAACCCCGCGNGGTGGTGAGCTAATTGCTTATCTC  427

Qy   81  GlyHisTyrTrpArgGlnAsnLeuValAlaAspGlyLeuLeuProLysGlyCysPro  100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  428  GGACATTACTGGCGTCAGAATCTGGTGGCAGATGGGCTGCTGCCAAAAGGCTGCCCG  487

Qy  101  GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  488  CAGTCTGGTCAGGTCGCCATTATTGCTGACGTAGATGAGCGTACCCGTAAAACAGGCGAA  547

Qy  121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr  140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  548  GCCTTCGCCGCCGGTCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG  607

Qy  141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla  160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  608  TCCAGTCCCGATCCGTTATTTAATCCCTTAAAAACCGGTGTTTGCCAACTGGATAACGCG  667

Qy  161  AsnValThrAspAlaIleLeuGluArgAlaGlyGlySerIleAlaAspPheThrGlyHis  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  668  AACGTGACTGACGCGATCCTTGAGCGCGCCGGAGGTTCAATTGCTGACTTTACGGGCCAT  727

Qy  181  TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys  200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  728  TATCAGACGGCCTTTCGCGAACTGGAACGGGTGCTTAATTTCCCGCAATCAAACTTGTGC  787

Qy  201  LeuLysArgGlyLysGlyAspAsnGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  788  CTTAAACGTGGAGAAGGACAGGAACGTGTTCATTAACGGCACGGCAATTACCATCGGAACTC  847

Qy  221  LysValSerAlaAspCysSerValSerLeuThrGlyAlaValAlaValSerLeuMetLeuThr  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  848  AAGGTGAGCGCGGACTGCAGCGTCTCACTAACCGGTGCGGTAGCCGTAAGCCTCGCCATCAATGCTGACG  907

Qy  241  GluIlePheLeuLeuGlnIleGlnAlaGlnGlyMetProGluProGlyTyrTrpGlyArgIleThr  260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  908  GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGTATTGGGGAAGGATCACC  967

Qy  261  AspSerHisGlnThrProAsnThrLeuLeuSerLeuHisAlaAsnAlaGlnPheAspLeuLeuGln  280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  968  GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGGGCAATTTATTTGTTTGCTACAA  1027

Qy  281  ArgThrProGluValAlaLysArgSerArgAlaAlaThrProLeuLeuAspLeuIleLysThrAla  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1028  CGCACACCCCAGAGGTTGCCCGGAGCCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG  1087

Qy  301  LeuThrProHisProProGlyLysGlnAlaThrTyrGlyValThrLeuProThrSerValLeu  320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1088  TTGACGCCCCATCCACCGGCAAAAACAGGCGTATGGGTGACATTACCCACTTCAGTGCTG  1147

Qy  321  PheIleAlaGlyHisAlaAspThrAsnLeuAlaAsnLeuGlyTyrGlyAlaLeuGluLeuAsnTrp  340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1148  TTTATCGCCGGACACGCAGATACTAATCTGGCAAATCTCGGGCGCGCACTCGGAGCTCAACTGG  1207

Qy  341  ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1208  ACGCTTCCCGGTCAGCCGGATAACACGCCGCCGGGTGGCGAACTGGTGTTTGAACGCTGG  1267

Qy  361  ArgArgLeuSerAspAlaAsnSerGlnThrTyrIleGlnValSerLeuValPheGlnThrLeuGln  380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1268  CGTCGGCTAAGCGATGCAAACAGCCAGACGTATATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG  1327

Qy  381  GlnMetArgAspGlyThrProLeuSerLeuAsnThrProProGlyGluValAlaLysLeuThr  400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1328  CAGATGCGTGATAAAAACGCCCCTGTCATTAAATAACCCCGCGGAGGTGGCAAAACTGACC  1387

Qy  401  LeuAlaGlyIleCysTyrGlyGluGluGlnLysAsnAlaLeuAlaGlyGlyMetCysSerSerLeu  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1388  CTGGCAGGAGTCGAAGAGCGGAAATGGCCAGGCATGTGTTCCTGTTGGCAGGGTTTTACGGCA  1447

Qy  421  IleValAlaAsnGluAlaArgIleProAlaCysSerSerLeu  432
         ||||||||||||||||||||||||||||||||||||||||||
Db 1448  ATCGTGAATGAAGCACGCCATACCGGCCGTGCAGTTTG  1483
```

```
Db   1208  ACGGTTCCCGTCAGCCGGATAACACGCCCGGATAACGCGTGGTGAACGCTGG 1267
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    361  ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380

Db   1268  CGTCGGCTAGGCGATAACAGCCAGTGGATTCAGGTTTCCGCTGGTCTTCCAGACTTTACAG 1327
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProLeuProGlyGluValLysLeuTr 400

Db   1328  CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCTGCCCGGCGAAGTGAAACTGAC 1387
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    401  LeuAlaGlyGluValLysLeuThrGlyGlyMetCysSerSerLeuAlaGlyPheThrGln 420

Db   1388  CTGGCAGGATGTGAAGACGGAAAATGCCAGGGCATGTGGTTTCGTTGGCAGGTTTTACGCAA 1447
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    421  IleValAsnGluIleArgIleProAlaCysSerLeu 432

Db   1448  ATCCTGAATGAAGCACGCATACCCGCGTGCCAGTTTG 1483
            |||||||||||||||||||||||||||||||||||||
Qy    1448  ATCCTGAATGAAGCACGCATACCCGCGTGCCAGTTTG 1483
```

RESULT 5

LOCUS   ECOAPPAA   1901 bp   DNA   linear   BCT 04-APR-2002
DEFINITION Escherichia coli periplasmic phosphoanhydride phosphohydrolase
            (appA) gene, complete cds.
ACCESSION   M58708
VERSION   M58708.1 GI:145283
KEYWORDS   .
SOURCE   Escherichia coli
  ORGANISM Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1 (bases 1 to 1901)
  AUTHORS   Dassa,J., Marck,C. and Boquet,P.L.
  TITLE   The complete nucleotide sequence of the Escherichia coli gene appA
            reveals significant homology between pH 2.5 acid phosphatase and
            glucose-1-phosphatase
  JOURNAL   J. Bacteriol. 172 (9), 5497-5500 (1990)
   PUBMED   2168385
REFERENCE   2 (sites)
  AUTHORS   Ostanin,K., Harms,E.H., Stevis,P.E., Kuciel,R., Zhou,M.M. and Van
            Etten,R.L.
  TITLE   Overexpression, site-directed mutagenesis, and mechanism of
            Escherichia coli acid phosphatase
  JOURNAL   J. Biol. Chem. 267 (32), 22830-22836 (1992)
   PUBMED   1429631
REFERENCE   3 (sites)
  AUTHORS   Ostanin,K. and Van Etten,R.L.
  TITLE   Asp304 of Escherichia coli acid phosphatase is involved in leaving
            group protonation
  JOURNAL   J. Biol. Chem. 268 (28), 20778-20784 (1993)
   PUBMED   8407904
FEATURES   Location/Qualifiers
     source   1..1901
                     /organism="Escherichia coli"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:562"
     gene   3..1522
                     /gene="appA"
     RBS   3..8
                     /gene="appA"
     CDS   16..108
                     /gene="appA"
                     /note="ORFX"
                     /codon_start=1
                     /transl_table=11
                     /product="unknown"
                     /protein_id="AAA72085.1"
                     /db_xref="GI:145284"
                     /translation="MMYLLWFVGILLMCSLSTLVLVWLLDPRLKS"
     stem_loop   118..149
                     /gene="appA"
     -35_signal   133..138
                     /gene="appA"
     -10_signal   156..161
                     /gene="appA"
     RBS   172..177
                     /gene="appA"
     CDS   188..1486
                     /gene="appA"
                     /codon_start=1
                     /transl_table=11
                     /product="acid phosphatase"
                     /protein_id="AAA72086.1"
                     /db_xref="GI:145285"
                     /translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP
                     TKATQLMQDVTPDAWPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQS
                     GQVAIIADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA
                     NVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPS
                     ELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRRITLDSHQWNTLLSLHNAQF
                     YLLQRTPEVARSRATPLLDLIKTALTPHPPQKQAYGVGVLPTSVLFIAGHDTNLANLGG
                     ALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNT
                     PGGEVKLTILAGCEERNAQGMCSLAGFTQIVNEARIPACSL"
     sig_peptide   188..253
                     /gene="appA"
     mat_peptide   254..1483
                     /gene="appA"
                     /product="periplasmic phosphoanhydride phosphohydrolase"
     variation   299..301
                     /gene="appA"
                     /standard_name="H303A mutant"
                     /note="created by site directed mutagenesis"
                     /citation=[2]
     variation   299..301
                     /gene="appA"
                     /standard_name="R16A mutant"
                     /note="created by site directed mutagenesis"
                     /citation=[2]
     variation   299..301
                     /gene="appA"
                     /standard_name="R20A mutant"
                     /note="created by site directed mutagenesis"
                     /citation=[2]
     variation   299..301
                     /gene="appA"
                     /standard_name="R92A mutant"
                     /note="created by site directed mutagenesis"
                     /citation=[2]
     variation   302..304
                     /gene="appA"
                     /standard_name="H17N mutant"
                     /note="created by site directed mutagenesis"
                     /citation=[2]
     variation   1163..1165
                     /gene="appA"
                     /standard_name="D304A mutant"
                     /note="created by site directed mutagenesis"
                     /citation=[3]
     variation   1163..1165
                     /gene="appA"
                     /phenotype="Asp 304 changed to Ala"
                     /replace="gct"
     terminator   1497..1522
                     /gene="appA"
                     /phenotype="Asp 304 changed to Glu"
                     /replace="gag"
                     /note="rho-independent terminator"
ORIGIN

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:

3.74e-196   Length: 1901
2182.00   Matches: 424
98.1%   Conservative: 0
```

```
Best Local Similarity:    98.1%              Mismatches:  8
Query Match:              96.6%              Indels:      0
DB:                       15                 Gaps:        0

US-10-601-319-10 (1-432) x ECOAPPAA (1-1901)

Qy      1  MetLysAlaIleLeuIleLeuProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla  20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188  ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTTATTCCGTTAACCCCGCAATCTGCA  247

Qy     21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly  40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248  TTCGCGCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT  307

Qy     41  ValAlaArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308  GTGGCGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA  367

Qy     61  ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGluLeuGluLeuIleAlaTyrLeu  80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    368  ACCTGGCCGGTGAAACTGGGTTGGCTGACACCGCGNGGTGAGCTGGAGCTAATCGCCTATCTC  427

Qy     81  GlyHisTyrTrpProArgGlnArgLeuValAlaAlaAspGlyLeuAlaValAlaAlaAspGlyCysGlyCysPro  100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428  GGACATTACCAAACGCCAGCGTCTGGGTGGCGGCAGATGGCGGTGGAAAAGGGCTGCCCG  487

Qy    101  GlnSerGlyInValAlaIleIleLeuAlaAspGlyLysThrArgLysThrArgGlyGlu  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488  CAGTCTGGTCAGGTCGGGATTATTGCTGATGGGAAGACGAGACGGTACCCGTAAAACAGGCGAA  547

Qy    121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleLeuThrValHisThrGlnAlaAspThr  140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548  GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG  607

Qy    141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAspAsnAla  160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608  TCCAGTCCCGAtCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAAACTGGATAACGCG  667

Qy    161  AsnValThrAspAlaIleLeuLeuValArgAlaGlySerIleAlaAspAspPheAlaPheHisHis  180
           |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db    668  AACGTGACTGACGCGGATCCTCAGCAGGGCCAGGAGGGTCAATTGCTGACTTTACCGGGCCAT  727

Qy    181  TyrGlnThrAlaPheArgGlyLeuGluAsnPheProGlnSerAsnLeuCys  200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    728  CGGCAAACGGCCTTTCGCGGACTGGAACGGGTTCTTAATTTCCGCAAATCAAACTTGTGC  787

Qy    201  LeuLysArgGlyLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGlyLeu  220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    788  CTTAAACGTGGAGAAACAGGACGAGAGCTGTTCATTAACGCAGGCCATTACCATCGGAACTC  847

Qy    221  LysValSerAlaAspCysSerValSerLeuThrGlyValAlaValSerLeuAlaSerMetLeuThr  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    848  AAGGTGAGCGCCGACAAATGTCTGTTAGGTGTCGGTGTAAGCCTCGGCATCAAATGCTGACG  907

Qy    241  GluIleProGluLeuGlnIleGluAlaGlnInAlaGlyMetProGlyGlyLeuTrpGlyIleThr  260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908  GAGATAATTCCTCTGCAACAACCACACAGGGAATGCCGGGGTGGGAAGGGGATACC  967

Qy    261  AspSerHisGlnIleTrpAspAsnThrLeuSerLeuHisAsnAlaGlyInPheAspLeuGln  280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968  GATTCACACCAGTGGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTGATTTATTGCTACAA  1027

Qy    281  ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028  CGCACGCCGGAAGTTGCCGCCCGCAGCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG  1087

Qy    301  LeuThrProHisSerProHisProGlyLeuValThrLeuProThrSerValLeu  320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088  TTGACGCCCCATCCACCCCAAAAACAGGGCGTAGGGGTGGTGACATTACCCACTTCAGTCTC  1147

Qy    321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148  TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGGGCGCACTGGAGCTCAACTGG  1207
```

```
RESULT 6
AX042376
LOCUS       AX042376               3470 bp    DNA     linear   PAT 23-NOV-2000
DEFINITION  Sequence 5 from Patent WO00064247.
ACCESSION   AX042376
VERSION     AX042376.1  GI:11340994
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Forsberg,C.W., Golovan,S. and Phillips,J.P.
  TITLE     Transgenic animals expressing salivary proteins
  JOURNAL   Patent: WO 0064247-A 5 02-NOV-2000;
            UNIVERSITY OF GUELPH (CA)
FEATURES             Location/Qualifiers
     source          1..3470
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="R15/APPA transgene"
ORIGIN

Alignment Scores:
Pred. No.:
Score:                    2182.00        Length:        3470
Percent Similarity:       98.1%          Matches:       424
Best Local Similarity:    96.6%          Conservative:  0
Query Match:              2              Mismatches:    8
DB:                                      Indels:        0
                                         Gaps:          0

US-10-601-319-10 (1-432) x AX042376 (1-3470)

Qy      1  MetLysAlaIleLeuIleLeuProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla  20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1811  ATGAAAGCCATCTTAATCCCATTTTTATCTCTTCTTATTCCGTTAACCCCGCAATCTGCA  1870

Qy     21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly  40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1871  TTCGCGCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT  1930

Qy     41  ValAlaArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1931  GTGGCGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA  1990

Qy     61  ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGluLeuGluLeuIleAlaTyrLeu  80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1991  ACCTGGCCGGTGAAACTGGGTTGGCTGACACCGCGTGGTGAGCTGGAGCTAATCGCCTATCTC  2050

Qy     81  GlyHisTyrTrpProArgGlnArgLeuValAlaAlaAspGlyLeuAlaValAlaAlaAspGlyCysGlyCysPro  100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2051  GGACATTACCAAACGCCAGCGTCTGGGTGGCGGCAGATGGCGGTGGAAAAGGGCTGCCCG  2110
```

(middle alignment block)

```
Qy    341  ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerArgLeuValPheGlnThrLeuGln  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208  ACGCGTTCCCGGTCAGCCGGTCAGGGATAACACCCCGCCAGGTGGGTGAACTGGTGTTTGA  1267

Qy    361  ArgArgLeuSerAspAsnSerGlnThrIleProIleGluValSerLeuValPheGlnThrLeuGln  380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268  CGTTCGGCGTAAGGCGGATAACAGCAGCGACTGAATTCAGGTTCGCGTGGTCTTCCAGACTTACAG  1327

Qy    381  GlnMetArgArgAspProLysGlyThrProLeuSerLeuAsnThrProProGlyLeuValLysLysSerThr  400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328  CAGATGCGTAGAAACGCCCGGTGCTCATTAAATACGCCGCCCCCGGAGGGTTAAAACTGACC  1387

Qy    401  LeuAlaGlyCysGluGluuArgAspAsnAlaGlnInGlyMetCysSerSerLeuAlaGlyPheThrGln  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388  CTGGCAGGATGTGAAGAGCGAAATGCCCAGGCCATGTGTTCGTTCGTTGCGGGTTTTACCCAA  1447

Qy    421  IleLeuAsnGluValAlaArgIleProAlaCysSerSerLeu  432
           ||||||||||||||||||||||||||||||||||||||||||||
Db   1448  ATCGTGAATGAAGCACACGGATACCGGGCCTGCAGTTTG  1483
```

```
RESULT 7
LOCUS       AX042374        4060 bp      DNA        linear     PAT 23-NOV-2000
DEFINITION  Sequence 3 from Patent WO0064247.
ACCESSION   AX042374
VERSION     AX042374.1  GI:11340992
KEYWORDS    synthetic construct
SOURCE      synthetic construct
  ORGANISM  other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Forsberg,C.W., Golovan,S. and Phillips,J.P.
  TITLE     Transgenic animals expressing salivary proteins
  JOURNAL   Patent: WO 0064247-A 3 02-NOV-2000;
            UNIVERSITY OF GUELPH (CA)
FEATURES             Location/Qualifiers
     source          1..4060
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="R15/APPA + intron transgene"
ORIGIN
```

Alignment Scores:
Pred. No.:

Score:	1.14e-195
Percent Similarity:	98.1%
Best Local Similarity:	98.1%
Query Match:	96.6%

Length:	4060
Matches:	424
Conservative:	0
Mismatches:	8
Indels:	0
Gaps:	0

DB:

US-10-601-319-10 (1-432) x AX042374 (1-4060)

```
Qy      1 MetLysAlaIleLeuIlePheProPheLeuSerLeuLeuIleLeuProLeuThrProGlnSerAla 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1811 ATGAAAGCCATCTTAATCCCATTTTTATCTCTTCTTGATTCGTACCCGCCAAACTGCA 1870

Qy     21 PheAlaGlnSerGluProLeuGluLeuLysGluSerValValIleValAlaSerArgHisGly 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1871 TTCGCTCAGAGTGAGCCGGAGCTGAAAGCTGAAAGTGTGGTGATTGTCGTCAGTGCATGGT 1930

Qy     41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1931 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCCACGCCA 1990

Qy     61 ThrTrpProValLysGlyGluGluThrProArgGlyGlyGluLeuLeuIleAlaThrTyrLeu 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1991 ACCTGGCCGGTAAAACTGGGTTGGCTGACAACCGCCCGGGTGAGCTAATCGCTATCTC 2050

Qy     81 GlyHisTyrThrArgProLeuGlnArgLeuValAlaAspGlyLeuLeuProLysGlyCysPro 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2051 GGACATTACCAACCGCCAGGCGTCTGGTGCGAAAAAGGGCTGCCCG 2110

Qy    101 GlnSerGlyGlnValAlaIleIleIleAlaAspValAlaAspGluGlyThrArgGlyThrGlyGlu 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2111 CAGTCCGGTCAGGTCGCGATTATTGCCGACGAGGGTACCCGTAAAAACAGGCGAA 2170

Qy    121 AlaPheAlaAlaGlyTyrLeuAlaProAspAspCysAlaIleThrValHisThrGlnIleAlaAspThr 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2171 GCCTTCGCCGCCGGCTGGCCACTTGACGTGCAATAACCCTACAACGCCAGGGCAGATACG 2230

Qy    141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2231 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGGTGTGCCAGTTGGATAACGCG 2290

Qy    161 AsnValThrAspAlaIleLeuGluArgValLeuAlaGlyGlySerIleAlaAspPheThrGlyHis 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2291 AACGTGACTGACCGGCATCCTCAGCAGGGCGAGGGGTCAATTGCTGACTTTACCGGGCAT 2350

Qy    181 TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2351 CGGCAAACCGCGTTTCGCGAACTGGAACGGGTGCTTAATTTCCCGCAATCAAAACTTGTGC 2410

Qy    201 LeuLysArgGlyLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2411 CTTAAACGTGAGAAAACAGACGAAAGCTGTTCATTAACGCAGGCACTGCCAAGTGAACTC 2470
```

```
RESULT 8
AX042375

LOCUS       AX042375       6116 bp    DNA    linear    PAT 23-NOV-2000
DEFINITION  Sequence 4 from Patent WO0064247.
ACCESSION   AX042375
VERSION     AX042375.1  GI:11340093
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Forsberg,C.W., Golovan,S. and Phillips,J.P.
  TITLE     Transgenic animals expressing salivary proteins
  JOURNAL   Patent: WO 0064247-A 4 02-NOV-2000;
            UNIVERSITY OF GUELPH (CA)
FEATURES             Location/Qualifiers
     source          1..6116
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="R15/APPA plasmid with pBLCArt3 vector"
ORIGIN
```

Alignment Scores:
Pred. No.:
Score:

Pred. No.:	2.07e-195		
Score:	2182.00		
Percent Similarity:	98.1%	Length:	6116
Best Local Similarity:	98.1%	Matches:	424
Query Match:	96.6%	Conservative:	0
		Mismatches:	8
		Indels:	0

DB: US-10-601-319-10 (1-432) x AX042375 (1-6116)   Gaps: 0

Qy   221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2471 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 2530

Qy   241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2531 GAGATATTTCCCTGCAACAGGCACAGGGAATGCCGGAGCCCGGGTGGGGAAGGATCACC 2590

Qy   261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuGln 280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2591 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAAGCGCCAATTTATTTGCTACAA 2650

Qy   281 ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2651 CGCACGCCAGAGGTTGCCGCGCCACCCGTTATTGGATTTGATCAAGACAGCG 2710

Qy   301 LeuThrProHisProProGlnLysGlyAlaTyrGlyValThrLeuProThrSerValLeu 320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2711 TTGACGCCCATCCACCCGCAAAAAAGGCGTATGGGTGTGACATTACCCACTTCAGTGCTG 2770

Qy   321 PheIleAlaGlyHisAspThrAsnLeuLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2771 TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG 2830

Qy   341 ThrLeuProGlyGlnProAspAlaAspCysAlaIleLeuAlaAspValAspGlyArgTrp 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2831 ACGCTTCCCGGTCAGCCGGATGCAGATTGCGCAATTCTGGCAGATGTCGACGGGCGCTGG 2890

Qy   361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2891 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 2950

Qy   381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2951 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACCCCGCCGGGCGAGGTGAAACTGACC 3010

Qy   401 LeuAlaGlyLysGluValArgAsnAlaGlnGlyLysMetCysSerIleAlaGlyTyrGln 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3011 CTGGCAGGATGTAAGAGGCGAAATGGCCAGGGCATGTGTTCCGTTGGCCGGTTTTACCGA 3070

Qy   421 IleValAlaAsnGluAlaArgIleProAlaCysSerSerLeu 432
         ||||||||||||||||||||||||||||||||||||||||||
Db 3071 ATCGTGAATGAAGCACGCCATACCCGCTTGCAGTTG 3106
```

Db: 2   Gaps: 0

```
Qy   201 LeuLysArgGluLysGlnAspSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2411 CTTAAACGTGAGAAACAGGACAGTTGCTCATTAACGACCAGCCATTACCATCGGAGACTC 2470

Qy   221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2471 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 2530

Qy   241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2531 GAGATATTTCCCTGCAACAGGCACAGGGAATGCCGGAGCCCGGGTGGGGAAGGATCACC 2590

Qy   261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuGln 280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2591 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAAGCGCCAATTTATTTGCTACAA 2650

Qy   281 ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2651 CGCACGCCAGAGGTTGCCGCGCCACCCGTTATTGGATTTGATCAAGACAGCG 2710

Qy   301 LeuThrProHisProProGlnLysGlyAlaTyrGlyValThrLeuProThrSerValLeu 320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2711 TTGACGCCCATCCACCCGCAAAAAAGGCGTATGGGTGTGACATTACCCACTTCAGTGCTG 2770

Qy   321 PheIleAlaGlyHisAspThrAsnLeuLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2771 TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG 2830

Qy   341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
RESULT 9
LOCUS       AX042373                6708 bp    DNA     linear   PAT 23-NOV-2000
DEFINITION  Sequence 2 from Patent WO0064247.
ACCESSION   AX042373
VERSION     AX042373.1  GI:11340991
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Forsberg,C.W., Golovan,S. and Phillips,J.P.
  TITLE     Transgenic animals expressing salivary proteins
  JOURNAL   Patent: WO 0064247-A 02-NOV-2000;
            UNIVERSITY OF GUELPH (CA)
FEATURES             Location/Qualifiers
     source          1..6708
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="R15/APPA + intron plasmid with pBLCAT3 vector"
ORIGIN

Alignment Scores:
Pred. No.:
Score:                 2.37e-195
Score:                 2182.00
Percent Similarity:    98.1%         Matches:        6708
Percent Similarity:    98.1%         Conservative:    424
Best Local Similarity: 98.1%         Mismatches:        8
Query Match:           96.6%         Indels:            0
DB:                    2             Gaps:              0

US-10-601-319-10 (1-432) x AX042373 (1-6708)

Qy     1  MetLysAlaIleLeuLeuProPheLeuSerLeuLeuProLeuThrProGlnSerAla   20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1811  ATGAAAGCCATCTTAATCCCATTTTAATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 1870

Qy    21  PheAlaGlnSerGluProLeuLeuLysGluSerValValIleValSerArgHisGly   40
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1871  TTCGCTCAGAGTGAGCCGCTGGAAAGTGTGGTAGTGATTGTCAGTCGTCGTCATGGT   1930

Qy    41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaProThr  60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1931  GTGCGTGCGCCACCAAGGCCACCCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA  1990

Qy    61  ThrTrpProValLysGluGlyLeuLeuValAlaAspProArgGlyTyrGlyGluLeuGln  80
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1991  ACCTGGCCGGTAAAACTGGGTTGGCTGGTCGCAGACCCCGCGAGCTAATCGGCTATCTC  2050

Qy    81  GlyHisThrTyrTrpArgGlnLeuAspGlyLeuLeuProLysCysGlyCysPro       100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2051  GGACATTACCAACCGCAGGTTCGTGGCGGATTGCTGCCGAAAAAGGGCTGCCCC       2110

Qy   101  GlnSerGlyGlnValAlaIleLeuAlaAspValAspGluArgThrArgLysThrGlyGlu 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2111  CAGTCTGGTCAGGTCGCGATTATTGCTGACGTGGACGAGCGTACCCGTAAAACAGGCGAA 2170

Qy   121  AlaPheAlaAlaGlyLeuAlaProAspAspCysAlaIleIleThrAlaHisGlnAlaThr 140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2171  GCCTTCGCCGCCGGTCTGGCACCTGACGATTGCGCAATCATCACCGCACATCAGGCCACG 2230

Qy   141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2231  TCCAGTCCCGATCCGCTGTTTAATCCTCTAAAAACCGGGGTTTGCCAACTGGATAACGCG 2290

Qy   161  AsnValThrAspAlaIleLeuGluArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2291  AACGTGACTGACGCGATCCTCGAGCGGGCAGGGGGCAGTATCGCAGACTTTACCGGGCAT 2350

Qy   181  IleValAsnGluAlaAsnGlyIleProAlaCysSerLeu                      432
          ||||||||||||||||||||||||||||||||||||
Db  2351  CGGCAAACCGGCGGTTTCGGCAACCGGTTGCTAATTTCCCGAATCAAACTTGTGC      2410

Qy   361  ArgArgLeuSerAspAsnArgSerGlnTrpIleGlnValSerLeuValPheGluGlnThrLeuGln 380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2891  CGTCGGCTRAGGCGATAACAGCCAGCTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 2950

Qy   381  GlnMetArgAspAsnSerLeuAsnThrProLeuSerLeuAsnThrProProGlyGluValValLysLeuValThr 400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2951  CAGATGCGTGATAATAGCCTGAACACGCCGCTGTCATTAAATACGCCCCCCGGAGAGGTGGTTAAACTGGTGACC 3010

Qy   401  LeuAlaGlyCysGluGluArgAsnAlaGlyGlnMetCysSerSerLeuAlaGlyPheThrGln 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3011  CTGGCCAGGATGTGAAGAGCGTAATCGCAGGGCAGATGTGTTCGTTCGGCAGTTTTACGCAA 3070

Qy   421  IleValAsnGluAlaAsnGlyIleProAlaCysSerLeu                      432
          ||||||||||||||||||||||||||||||||||||
Db  3071  ATCGTGAATGAAGCACGCATACCCGCTTGCAGTTTG                         3106

Qy   341  ThrLeuProGlyGlnProAspAsnAsnThrProProGlyGlyValGluLeuValPheGluArgTrp 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2831  ACGGCTTCCCGGTCAGCCCGGCGATAACACGCCCAGGTGGGTGAAACTGGTGTTTGAACGCTGG 2890
```

```
Qy   201  LeuLysArgGluLysGlnAspGluSerCysSerSerLeuThrGlnAlaLeuProSerGluLeu 220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2411  CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTCAGGTCGCGCATTACCATCGGAACTC 2470

Qy   221  LysValSerAlaAspPheCysValValSerLeuThrGlyValAlaValSerLeuAlaSerMetLeuThr 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2471  AAGGTGAGCGCCGACAAAATGTCTCATTAACCGGTGGGGTAGCCGTTCCGCATCAAATGTCACG 2530

Qy   241  GluIlePheLeuLeuLeuGlnGlnIleAlaGlnGlyMetProGluGlyTrpGlyTyrArgIleThr 260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2531  GAGATATTTCTCCTGCAACAAGCACAGGAATGCCGGAGGCGGGTGGGGAAGGATCACC 2590

Qy   261  AspSerArgHisGlnThrAspAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2591  GATTCACACCAGTGGAAAACACCTTGCTAAGTTTGCAATAATCAGCCAATTTTATTGCTAACAA 2650

Qy   281  ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleLeuLysThrAla 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2651  CGCACGCCAGAGTTGCCCGCCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 2710

Qy   301  LeuThrProHisProProGlnLeuGlnAlaLeuTyrGlyValThrLeuProThrSerValLeu 320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2711  TTGACGCCCCATCCACCGCAAAAACAGGCGCTATGGGTGGTGACATTACCCACTTCAGTGCTG 2770

Qy   321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2771  TTTATCGCCGGGACACGAATACTAATCTGGCAAATCTCGGCGCACTGGAGCTCAACTGG 2830

Qy   361  ArgArgLeuSerAspAsnArgSerGlnTrpIleGlnValSerLeuValPheGluGlnThrLeuGln 380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2891  CGTCGGCTRAGGCGATAACAGCCAGCTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 2950

Qy   381  GlnMetArgAspAsnSerLeuAsnThrProLeuSerLeuAsnThrProProGlyGluValValLysLeuValThr 400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2951  CAGATGCGTGATAATAGCCTGAACACGCCGCTGTCATTAAATACGCCCCCCGGAGAGGTGGTTAAACTGGTGACC 3010

Qy   401  LeuAlaGlyCysGluGluArgAsnAlaGlyGlnMetCysSerSerLeuAlaGlyPheThrGln 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3011  CTGGCCAGGATGTGAAGAGCGTAATCGCAGGGCAGATGTGTTCGTTCGGCAGTTTTACGCAA 3070

Qy   421  IleValAsnGluAlaAsnGlyIleProAlaCysSerLeu                      432
          ||||||||||||||||||||||||||||||||||||
Db  3071  ATCGTGAATGAAGCACGCATACCCGCTTGCAGTTTG                         3106

RESULT 10
LOCUS       AX042378               17732 bp    DNA     linear   PAT 23-NOV-2000
DEFINITION  Sequence 7 from Patent WO0064247.
ACCESSION   AX042378
VERSION     AX042378.1  GI:11340996
```

Left column

```
KEYWORDS    synthetic construct
SOURCE      synthetic construct
ORGANISM    other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Forsberg,C.W., Golovan,S. and Phillips,J.P.
  TITLE     Transgenic animals expressing salivary proteins
  JOURNAL   Patent: WO 0064247-A 7 02-NOV-2000;
            UNIVERSITY OF GUELPH (CA)
FEATURES             Location/Qualifiers
     source          1..17732
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Lama2/APPA transgene"
ORIGIN

Alignment Scores:
Pred. No.:
Score:                    9.81e-195
                          2182.00
Percent Similarity:       98.1%
Best Local Similarity:    98.1%
Query Match:              96.6%

Length:          17732
Matches:           424
Conservative:        0
Mismatches:          8
Indels:              0
Gaps:                0
```

US-10-601-319-10 (1-432) x AX042378 (1-17732)

```
Qy     1  MetLysAlaIleLeuLeuIleLeuProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla   20
Db 12653  ATGAAAGCCATCTTAATCCCATTTTAATCTCTTCTGAATCCGTTAACCCGCAATCTGCA  12712

Qy    21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly   40
Db 12713  TTCGCTCAGAGTGAGCCCGAGCTGAAGCTGGAAAGTGTCGTGATTGTCAGTCGTCATGGT  12772

Qy    41  ValArgAlaProThrLysValThrGlnLeuMetGlnAspValThrProAspAlaTrpPro   60
Db 12773  GTGCGCGTCCAACCAAAGGCCACGCAACTGATGCAAGACGTCACCCCAGACGCCATGCCA  12832

Qy    61  ThrTrpProValGlyLeuGlyValGluThrProArgGlyGluLeuIleAlaTyrLeu   80
Db 12833  ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGCGGTGAGCTNATCGCCTATCTC  12892

Qy    81  GlyHisSerTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyPro  100
Db 12893  GGACATTACCAACGCCAGCGGCTGGTAGCCGACGGATTGCTCGGCGAAAAAGGCTCCCCG  12952

Qy   101  GlnSerGlyValAlaIleAlaIleLeuAlaAspValAspArgThrArgLysThrGlyGlu  120
Db 12953  CAGTCTCGGTCAGGTCCGGATTATTGCTGGACGTCGACGAGGCGTACCCGTAAAAACAGG  13012

Qy   121  AlaPheAlaLeuAlaGlyLeuGluArgValAlaGlyGlySerIleAlaAspPheThr  140
Db 13013  GCCTTCGCCCTGGCCGGCCTTGGCCACCTGACGTGCAATAACCGGCCAGGATACG  13072

Qy   141  SerSerProAspProLeuPheAsnProLeuLysThrGlyTyrValLysAspAsnAla  160
Db 13073  TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGACGTTTGCCAACTGGATAACGCG  13132

Qy   161  AsnValThrArgAspAlaIleLeuGluGlyAlaGlyLysSerIleAlaAspPheThrGlyHis  180
Db 13133  AACGTGACTCGAGACGGCGATCCTCAGCGAGGGGTCAAATGCTACATTTACCGGGCAT  13192

Qy   181  TyrGlnThrArgAlaPheArgGluLeuGluArgValAlaValLeuAsnPheProGlnSerAsnLeuCys  200
Db 13193  CGGCAAAACGGCGTTCCGGCAACTGGAACGGGGTGCTTAATTTCCGCAATCAAACTTGTGC  13252

Qy   201  LeuLysArgGlyLysGlnAspSerCysSerLeuThrArgGlnAlaLeuProSerGluLeu  220
Db 13253  CTTAAACGTGGACAAACAAGGCGTGTTCATTAACGACAGGCATTACCATCGGGACTC  13312

Qy   221  LysValSerAlaAspCysValSerLeuLeuThrGlyAlaValSerLeuAlaSerMetLeuThr  240
Db 13313  AAGGTGAGCGCCGACAAATGTCTCATTAACCGGTGCGCGGTAAGCCTCGCCATCAATGCTGACG  13372
```

Right column

```
Db 13373  GAGATATTTCTTCCTGCGAACAAGCACACGGGAAATGCCCGAGCCCGGGTGGGGGGAATCACC  13432

Qy   261  AspSerHisGlnTrpAspAsnThrLeuLeuSerLeuHisAlaAsnAlaGlnnPheAspLeuGln  280
Db 13433  GAATTCACACCAGTGGGAACAACCTTGCTAAGTTTCAATAACGCGCAAATTTATTTGCTACAA  13492

Qy   281  ArgThrProGluValAlaLeuAlaArgSerArgAlaThrProLeuLeuAspLeuLeuLysThrAla  300
Db 13493  CGCACACGCCAGAGGTTGCCCGCAGCCCGCGCCCACCCCGTTATTAGGATTTGAACAAGACAGCG  13552

Qy   301  LeuThrProHisProProGlnLysGlnIleAlaThrTyrGlyValThrLeuProThrSerValLeu  320
Db 13553  TTGACGCCCCATCCACCGCAAAAACAGGCGTNATGGGTGTGACAATTACCCACTTCAGTGCTG  13612

Qy   321  PheIleAlaGlyHisAspThrAsnLeuValAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340
Db 13613  TTTATACCGGCCATGACACGAATAACAGCCGTTCGGCTGGAGCTCAACTGG  13672

Qy   341  ThrLeuProGlyGlnProAspAsnThrProProGlyCysGlyLeuGluLeuValPheGluArgTrp  360
Db 13673  ACGCTTCCCGGTCAGCCCGGATAACACGCCCCGTGGTGGGTGAACTGGTGTTTGAAGCGTGG  13732

Qy   361  ArgArgLeuSerArgAspAspSerGlnTrpIleGluGlnValSerArgLeuValPheGlnThrLeuGln  380
Db 13733  CGTCGGCTAAGCGTAGGACGATAACAGCCCGTCAGTTTCGCTGGTTGTTCCAGACTTTACAG  13792

Qy   381  GlnMetArgAspLysThrProLeuSerLeuAsnAsnThrArgProGlyGluValAlaValLysLeuLeuThr  400
Db 13793  CAGATGCGTGATAAAACCCCGCTGTCATTAAATAACACGCGCCCGGGAGAGGTGGCGGTAAAACTGACC  13852

Qy   401  LeuAlaGlyTyrCysGluGluAlaArgGlnAsnAlaGlnnIleGlyMetCysSerLeuAlaGlyTyrPheThrGln  420
Db 13853  CTGGCAGGATATTGTGAAGAGGCGAAAATGCCCAGGCGCATGTGTTCCTTGGCAGGTTTTACCCAA  13912

Qy   421  IleValAsnGluValAlaArgIleLeuProAlaCysSerSerLeu  432
Db 13913  ATCGTGAATGAAGCACCGCATACCCGCTTGCAGTTTG  13948
```

```
RESULT 11
LOCUS       AX042372                20623 bp    DNA     linear   PAT 23-NOV-2000
DEFINITION  AX042372 Sequence 1 from Patent WO0064247.
ACCESSION   AX042372
VERSION     AX042372.1  GI:11340090
KEYWORDS    .
SOURCE      synthetic construct
ORGANISM    synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Forsberg,C.W., Golovan,S. and Phillips,J.P.
  TITLE     Transgenic animals expressing salivary proteins
  JOURNAL   Patent: WO 0064247-A 1 02-NOV-2000;
            UNIVERSITY OF GUELPH (CA)
FEATURES             Location/Qualifiers
     source          1..20623
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Lama2/APPA plasmid"
ORIGIN

Alignment Scores:
Pred. No.:
Score:                    1.22e-194
                          2182.00
Percent Similarity:       98.1%
Best Local Similarity:    98.1%
Query Match:              96.6%

Length:          20623
Matches:           424
Conservative:        0
Mismatches:          8
Indels:              0
Gaps:                0
```

US-10-601-319-10 (1-432) x AX043372 (1-20623)

```
Qy     1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla   20
Db 12653 ATGAAAGCCATCTTAATCCCATTTTTATCTCTGATTCCGTTAACCCCGCAATCTGCA 12712

Qy    21 PheAlaGlnSerGluProLeuLeuLysGluLeuSerValValIleValSerArgHisGly   40
Db 12713 TTCGCTCAGAGTGAGCCGGACTGAGCTGAAAGTGTGGTCATCGTCGTCATGGT 12772

Qy    41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro   60
Db 12773 GTTCGTGCGCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 12832

Qy    61 ThrTrpProValLysGlyGlyGluLeuThrProArgGlyGlyGlnLeuGluIleAlaTyrLeu   80
Db 12833 ACCTGGCCGGTGAAAGGTGGCGAACTGACGCCGCGTGGTGGTCAGCTGGAGATCGCCTATCTC 12892

Qy    81 GlyHisTyrTrpLysGlnLeuArgGlnLeuValAlaAlaAspGlyLeuLeuProLysCysGlyPro  100
Db 12893 GGACATTACCAACCCCAGGGTCTCGGTAGCCGACCGGATTCCTGGGAAAAAGGGGCTGCCCG 12952

Qy   101 GlnSerGlyLnValAlaIleIleIleAlaAspValAspGlyThrArgGlyLysThrGlyGlu  120
Db 12953 CAGTCTGGTCAGTTGCGATTATTGCGGATGTCGACGGAAGCCGTAAAAACAGGGCGAA 13012

Qy   121 AlaPheAlaAlaGlyLeuAlaProAspCysGlyAlaIleThrValHisSerThrGlnAlaAspThr  140
Db 13013 GCCTTCGCCGCCGGTCTGGCGCCTGACTGTGGCAATAACCGTACATACAGGCAGATACG 13072

Qy   141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAspAsnAla  160
Db 13073 TCCAGTCCCGATCCGTTATTAATCCCTCTAAAAACTGGCGTGTGTCCAACTGGATAACCCG 13132

Qy   161 AsnValThrAspAlaIleLeuLeuGluArgGlyAlaLeuGlySerIleLeuAlaAspPheThrGlyHis  180
Db 13133 AACGTGACTGACGCGATCCTCAGCGAGGGTGCATTAGGCTCAATTTTAGCTTTACCGGGGCAT 13192

Qy   181 TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheAsnProGlnSerAsnLeuCys  200
Db 13193 CGGCAAACGGCGTTTCGGGAACTGGAACGGGTGCTGAATTTTAATCCGCAAAGCAACCTGTGC 13252

Qy   201 LeuLysArgGlyLysGluGlnTrpAspAspSerCysSerSerLeuThrGlnAlaLeuProSerGluGlu  220
Db 13253 CTTAAACGTGGAAAACAGGAAAGCTGTTCATTAACCCAGGCATTACCATCGGAACTC 13312

Qy   221 LysValSerAlaAlaAspCysValValSerLeuThrArgGlyAlaValAlaSerMetLeuThr  240
Db 13313 AAGGTGAGCGCCGCCGACAATGTCTCATTAACCGCGTAAGCCTCGCCATCAATGCTGACG 13372

Qy   241 GluIlePheLeuLeuGlnIleGlnAsnGlyMetProGluProGlyTrpGlyTyrArgIleThr  260
Db 13373 GAGATATTTCTCCTTGCAACAGCAAACCAGGCGTATGGGTGACATTAGACTGGGAAGGAGATCACC 13432

Qy   261 AspSerHisGlnTrpAspAsnThrLeuLeuHisAlaAsnLeuAlaGlnPheAspAspLeuGln  280
Db 13433 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATGCGCAATTTATTATTGCTACGA 13492

Qy   281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspProLeuIleLeuLysThrAla  300
Db 13493 CGCACGCCAGAGGTTGCCCGCAGCCGCGCGACCCCGTTATTGGATCCGTTAAGAACAGCG 13552

Qy   301 LeuThrProHisProGlnLeuSerGlnAlaTyrGlyValThrLeuProThrThrSerValLeu  320
Db 13553 TTGACGCCCCATCCACAGCAAAACAGCGTATGGGTGTGACATTACCACTTCAGTGCTG 13612

Qy   321 PheIleLeuAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340
Db 13613 TTTATCCGCCGACACGATACTAATCTGGCGAAACCTCCGGCGCCACTCGGAACTGAACTGG 13672

Qy   341 ThrLeuProGlyGlnProGlnAspProAsnThrProProGlyGlyLeuGluLeuValPheGluArgTrp  360
Db 13673 ACGCTTCCCGGTCAGCCGCAGGATCCGGATAAACACGCCGCCGGAGGCCTTGAACTGG 13732

Qy   361 ArgArgLeuSerArgAspAsnSerGlnAsnSerGlnThrTrpIleGluGlnIleValSerLeuValLeuPheGlnAsnThrLeuGln  380
Db 13733 CGTCGGCTAAGGCGATAAACAGCCAGGGAGGATAACAGGCCAGGGAATTCAGGTTCGCCTGGTGCTTCCAGACTTTACAG 13792

Qy   381 GlnMetArgAsnProLysThrProLeuSerLeuAsnThrProAsnProGlyProGlyTyrGlyGluValValLysLeuThr  400
Db 13793 CAGATGCGTGATAAAAACGCCGGCTGTCGATTAAATACCGCCGCCGGAGGGTGCTTCCAGACTTTGACC 13852

Qy   401 LeuAlaGlyCysGluGluArgGlnAsnAlaGlyIleGlyMetCysSerArgLeuAlaGlyLyPheThrGln  420
Db 13853 CTGGCAGGATGCGAAGAGCGAAAATCCGCACAGGGCCGTGTCGGCAGGTTTTTACGCAA 13912

Qy   421 IleValAsnGluValAlaArgIleProAlaCysSerLeu  432
Db 13913 ATCGTGAATGAAGCACGGCATACCCCCGCTTGCAGTTTG 13948
```

RESULT 12
U00096_10
WPCOMMENT
Sequence split into 47 fragments   LOCUS U00096   Accession U00096

Fragment Name	Begin	End
U00096_00	1	110000
U00096_01	100001	210000
U00096_02	200001	310000
U00096_03	300001	410000
U00096_04	400001	510000
U00096_05	500001	610000
U00096_06	600001	710000
U00096_07	700001	810000
U00096_08	800001	910000
U00096_09	900001	1010000
U00096_10	1000001	1110000
U00096_11	1100001	1210000
U00096_12	1200001	1310000
U00096_13	1300001	1410000
U00096_14	1400001	1510000
U00096_15	1500001	1610000
U00096_16	1600001	1710000
U00096_17	1700001	1810000
U00096_18	1800001	1910000
U00096_19	1900001	2010000
U00096_20	2000001	2110000
U00096_21	2100001	2210000
U00096_22	2200001	2310000
U00096_23	2300001	2410000
U00096_24	2400001	2510000
U00096_25	2500001	2610000
U00096_26	2600001	2710000
U00096_27	2700001	2810000
U00096_28	2800001	2910000
U00096_29	2900001	3010000
U00096_30	3000001	3110000
U00096_31	3100001	3210000
U00096_32	3200001	3310000
U00096_33	3300001	3410000
U00096_34	3400001	3510000
U00096_35	3500001	3610000
U00096_36	3600001	3710000
U00096_37	3700001	3810000
U00096_38	3800001	3910000
U00096_39	3900001	4010000
U00096_40	4000001	4110000
U00096_41	4100001	4210000
U00096_42	4200001	4310000
U00096_43	4300001	4410000
U00096_44	4400001	4510000
U00096_45	4500001	4610000
U00096_46	4600001	4635675

Continuation (11 of 47) of U00096 from base 1000001 (U00096 Escherichia coli K-12 MG1655

Alignment Scores:
Pred. No. :   1.42e-193   Length: 110000

```
Score:                2182.00        Matches:         424
Percent Similarity:   98.1%          Conservative:      0
Best Local Similarity: 98.1%         Mismatches:        8
Query Match:          96.6%          Indels:            0
DB:                   15             Gaps:              0

US-10-601-319-10 (1-432) x U00096_10 (1-1100000)

Qy     1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla    20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 39840 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTCGATTCCCTTAACCCCCAAATCTGCA 39899

Qy    21 PheAlaGlnSerGluProGluLeuLysLeuSerValValIleValSerArgHisGly      40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 39900 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGTCGGTCGTGATTGTCAGTCGTCATGGT     39959

Qy    41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro    60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 39960 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 40019

Qy    61 ThrTrpProValLysLeuGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu    80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40020 ACCTGGCCGGTAAAACTGGGTGAACTCACGCCGCGTGGTGGCGAGTTAATCGCCTATCTC 40079

Qy    81 GlyHisTyrTrpArgGlnArgLeuValAlaAlaGlyLeuLeuProLysCysGlyLysPro   100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40080 GGACATTACCAACGCCAGCGTCTGGTAGCCGCAGGCGGATTGCTGCCGAAAAAGGGCTGCCCG 40139

Qy   101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu   120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40140 CAGTCTGGTCAGGTCGCGATTATTGCTGACGTCGACGAGCGTACCCGTAAAACAGGCGAA 40199

Qy   121 AlaPheAlaAlaGlyLeuAlaAlaProAspCysByalAlaIleThrValHisThrGlnAlaAspThr 140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40200 GCCTTCGCCGCCGGCCTGGCCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 40259

Qy   141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla   160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40260 TCCAGTCCCGATCCGTTATTTAATCCCTAAAAACTGGCGTTGCCAACTGGATAACGCG 40319

Qy   161 AsnValThrAspAlaIleLeuGluValArgValLeuAsnPheProGlnSerAsnLeuCys   180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40320 AACGTGACTGACGCGATCCTCGAACGGGCGGTCAAATGGCTTAAATTTCCGCAGAGCCAT 40379

Qy   181 TyrGlnThrLeuAlaPheArgGluLeuGlyValLeuArgValLeuAsnPheProThrGlyHis 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40380 CGGCAAACGGGCGTTTCCGGAACTGGAAGGGCGTTGCTTAATTTCCGCAATCAAAACTTGTGC 40439

Qy   201 LeuLysArgGluLysLeuAspGluSerCysSerSerLeuThrGlnIleAlaLeuProSerGluLeu 220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40440 CTTAAACGTGAGAAAACAGGACGAAAGTGTTCATTAACGGACATCGGAACTC 40499

Qy   221 LysValSerArgAlaAspCysValSerValSerLeuThrArgGlyAlaValSerLeuAlaSerMetLeuThr 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40500 AAGGTGAGCGCGGACAAATGTCTCATTAACCGGTGCCGGTAAGCCTCGCCATCAAATGCTGACG 40559

Qy   241 GluIleLeuPheLeuLeuGlyLeuAsnGlyLeuMetProGlyProGlyTyrTrpGlyIleThr 260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40560 GAGATATTTCCTGCAACAGACACCTTGCTAAGTTGCATAACGCCGGGGGAGGGAGGATCACC 40619

Qy   261 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuLeuIleLysThrAla 280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40620 GATTCACACCAGTGGAACACCCTTGCTAGTTGCATAACGCCCGGGGAGGGAGGATCACC 40679

Qy   281 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40680 CGCACCCCAGAGGTTGCCCGCAGCCCGCCTTATTGAGATTTGATCAAGACAGCG 40739

Qy   301 LeuThrProHisProProGlnIleLysGlnIleAlaThrTyrGlyValThrLeuProThrSerValLeu 320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40740 TTGACGCCCCATCCACCGCAAAAACAGGGCGTTATGGGTGACATTACCACTTCAGTCCTG 40799

Qy   321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaAlaLeuGluLeuAsnTrp 340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40800 TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGGCGGCCACTGGAGCTCAACTGG 40859

Qy   341 ThrLeuProGlyGlnProAspAsnAsnThrProProProGlyLeuGlyLeuGluLeuValPheGlnThrLeuGln 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40860 ACGCTTCCCGGTTCAGCCGGGATAACACCGCCGGATAACCACGGTGGTGGTGAACTGTGG 40919

Qy   361 ArgArgLeuSerArgAspAspAsnSerGlnThrTrpIleGluGlnIleValSerIleLeuValPheGlnThrLeuGln 380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40920 CGTCGGCGGCCTTAAGCCGGTATAACAGCCAGTGGATTGAGCAGATTGTTTCGATTCTGGTGTTTCCAGACTTTACAG 40979

Qy   381 GlnMetArgGlyAspGlyThrProLeuLeuProGlyLeuSerLeuAsnThrProProProGlyLeuGly...GluThr 400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 40980 CAGATGCGTGGTGATAAAAACGCCCGTGTGTCATTAAATACGCCGCCCCCGGGAGAGGTGAAAACTGACC 41039

Qy   401 LeuAlaGlyLeuCysGluGlyValArgAsnAlaLeuGlnGlyLeuGlyMetCysSerSerLeuAlaGlyPheThrGln 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 41040 CTGGCAGGATGTTGAAGGAGCGGAAATGCCCAGGGCGATGTGTTCGTTCGTTGGGCAGGGTTTTTTACGGCAA 41099

Qy   421 IleValAlaAsnGluValAlaArgIleLeuProAlaAlaCysSerLeu 432
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 41100 ATCGTGAATGAAGCACCCATACCGGGCGCAGTTG 41135
```

RESULT 13
AP009048_10
WPCOMMENT
Sequence split into 47 fragments   LOCUS AP009048   Accession AP009048

Fragment Name	Begin	End
AP009048_00	1	110000
AP009048_01	100001	210000
AP009048_02	200001	310000
AP009048_03	300001	410000
AP009048_04	400001	510000
AP009048_05	500001	610000
AP009048_06	600001	710000
AP009048_07	700001	810000
AP009048_08	800001	910000
AP009048_09	900001	1010000
AP009048_10	1000001	1110000
AP009048_11	1100001	1210000
AP009048_12	1200001	1310000
AP009048_13	1300001	1410000
AP009048_14	1400001	1510000
AP009048_15	1500001	1610000
AP009048_16	1600001	1710000
AP009048_17	1700001	1810000
AP009048_18	1800001	1910000
AP009048_19	1900001	2010000
AP009048_20	2000001	2110000
AP009048_21	2100001	2210000
AP009048_22	2200001	2310000
AP009048_23	2300001	2410000
AP009048_24	2400001	2510000
AP009048_25	2500001	2610000
AP009048_26	2600001	2710000
AP009048_27	2700001	2810000
AP009048_28	2800001	2910000
AP009048_29	2900001	3010000
AP009048_30	3000001	3110000
AP009048_31	3100001	3210000
AP009048_32	3200001	3310000
AP009048_33	3300001	3410000
AP009048_34	3400001	3510000
AP009048_35	3500001	3610000
AP009048_36	3600001	3710000
AP009048_37	3700001	3810000
AP009048_38	3800001	3910000
AP009048_39	3900001	4010000
AP009048_40	4000001	4110000
AP009048_41	4100001	4210000
AP009048_42	4200001	4310000
AP009048_43	4300001	4410000
AP009048_44	4400001	4510000

Alignment Scores:

AP009048_45	4500001	4610000
AP009048_46	4600001	4646332
Continuation [11 of 47] of AP009048 from base 1000001 (AP009048 Escherichia coli W3110 D

Pred. No.:
Score: 1.42e-193
Score: 2182.00
Percent Similarity: 98.1%
Percent Similarity: 98.1%
Best Local Similarity: 98.6%
Query Match: 96.6%
DB: 15

Length: 110000
Matches: 424
Mismatches: 0
Conservative: 0
Mismatches: 8
Indels: 0
Query Gaps: 0
Gaps: 0

US-10-601-319-10 (1-432) x AP009048_10 (1-110000)

```
Qy    1 MetLysAlaIleLeuIleLeuProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 41039 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATCCCGTTAACCCCGCAATCTGCA 41098

Qy   21 PheAlaGlnSerGluProGlyLeuLysLeuGluSerValValIleValSerArgHisGly 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 41099 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATCGTCAGTCGTCATGGT 41158

Qy   41 ValArgGlyAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 41159 GTGCGTGGTGCTCCAACCAAGGCCACACAGCTGATGCAGGATGTCACCCCAGACGCATGGCCA 41218

Qy   61 ThrTrpProValGlyLeuGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 41219 ACCTGGCCGGTAAAACTGGGTTGGCGTGACACCGGCGGTGGAGCTTATCGCCTTATCTC 41278

Qy   81 GlyHisLeuTyrThrArgGlnAsnArgLeuValAlaAlaGlyLeuLeuProLysCysGlyCysPro 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 41279 GGACATTACCAACGCCAGCGTCTGTAGCCGACCGGATTCCTGCCGGCAAAAAGGGCTGCCCG 41338

Qy  101 GlnSerGlyIleValAlaIleAlaIleLeuAlaAspValAlaAspGluArgThrArgGlyThrGlyGlu 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 41339 CAGTCTCGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTGGCACAGGCGAA 41398

Qy  121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 41399 GCCTTCGCGCCCGGCCTGGCACTGACCTGCATAACACGTGCACTACCCAGGCGACATACG 41458

Qy  141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAspAsnAla 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 41459 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACCTGGCGTTTGCCAACTGGATAACCCG 41518

Qy  161 AsnValThrAspAlaIleLeuLeuGluValArgAspArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 41519 AACGTGACTGACGCGATCCTCAGCGAGGGTCAAATGCCTAGACTTTACCGGGCAT 41578

Qy  181 TyrGlnThrAlaPheArgGluLeuUrgAspLeuGlySerAsnPheProGlnSerAsnLeuCys 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 41579 CGGCAAACGGCGTTTCGGGAACTGGACAGGGGGTTCAATTGCCAATCAAAACCTTGTGC 41638

Qy  201 LeuLysArgGlyLysGlyLysAspCysGluSerLeuThrArgGlnAlaLeuProSerGlu 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 41639 CTTAAACGTGAGAAACAGGACGAAAGCGTGTTCATTACCAGGCATACCATGGGAACTC 41698

Qy  221 LysValSerAlaAspCysValSerValSerLeuAlaSerMetLeuThr 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 41699 AAGGTGAGCGGCCGACAATGTCTCATTAACCGGTGGCCTCGCATCAATGCTGACG 41758

Qy  241 GluIlePheLeuLeuGlnInAlaGlyIleGlyMetProGlyTrpGlyTyrArgIleThr 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 41759 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGGGGTGGGGTATCGGATCACC 41818

Qy  261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 41819 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTATTGCTACAA 41878

Qy  281 ArgThrProGluValAlaAsrSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 41879 CGCACGCCAGAGGTTGCCGCGCCACCCGTTATTAGATTTGATCAAGACAGCG 41938

Qy  301 LeuThrProHisProProGlnInLysGlnInAlaTyrGlyValThrLeuProThrSerVal 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 41939 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCGACTTCAGTGCTG 41998

Qy  321 PheIleAlaGlyTyrHisAspThrAsnLeuAlaAsnLeuGlyLysGlyValThrLeuProThrSerVal 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 41999 TTTATCGCCGGACACGATACTAATCTGGCAAAATCTCGGGCGGCGGCGGCCACTGAACTGG 42058

Qy  341 ThrLeuProGlyGlnInProAspAsnThrProProGlyGlyGluLeuGlyArgTrp 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 42059 ACGCTTCCCGGTCAGCCGGATAACACGCCCCGGTGGTGAACTGGTGTTGAACCGTGG 42118

Qy  361 ArgArgLeuSerArgAspAsnSerGlnInThrProIleGluIleValSerValPheGlnInThrLeuGln 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 42119 CGTCGGCTAAGCGGATAAACAGCCAGTGGAATCAGGTTTCGGTTCCAGATTTACAG 42178

Qy  381 GlnMetArgAspProLeuTyrThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 42179 CAGATGCGTGATAAAAACGCCCTGTACACTCCATTAAATACGGCGAGGAGGTGAAACTGACC 42238

Qy  401 LeuAlaGlyCysGluGluArgAsnAlaGlnInMetCysSerLeuAlaGlyPheThrGln 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 42239 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGTCATGTGTCGTTGGCAGGTTTTACGCAA 42298

Qy  421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 42299 ATCGTTGAAATGAAGCACGCATACCGGCCGGTGCAGTTTG 42334
```

RESULT 14
CP000038_10
WPCOMMENT
Sequence split into 49 fragments   LOCUS CP000038 Accession CP000038

Fragment Name	Begin	End
CP000038_00	1	110000
CP000038_01	100001	210000
CP000038_02	200001	310000
CP000038_03	300001	410000
CP000038_04	400001	510000
CP000038_05	500001	610000
CP000038_06	600001	710000
CP000038_07	700001	810000
CP000038_08	800001	910000
CP000038_09	900001	1010000
CP000038_10	1000001	1110000
CP000038_11	1100001	1210000
CP000038_12	1200001	1310000
CP000038_13	1300001	1410000
CP000038_14	1400001	1510000
CP000038_15	1500001	1610000
CP000038_16	1600001	1710000
CP000038_17	1700001	1810000
CP000038_18	1800001	1910000
CP000038_19	1900001	2010000
CP000038_20	2000001	2110000
CP000038_21	2100001	2210000
CP000038_22	2200001	2310000
CP000038_23	2300001	2410000
CP000038_24	2400001	2510000
CP000038_25	2500001	2610000
CP000038_26	2600001	2710000
CP000038_27	2700001	2810000
CP000038_28	2800001	2910000
CP000038_29	2900001	3010000
CP000038_30	3000001	3110000
CP000038_31	3100001	3210000
CP000038_32	3200001	3310000
CP000038_33	3300001	3410000
CP000038_34	3400001	3510000
CP000038_35	3500001	3610000
CP000038_36	3600001	3710000
CP000038_37	3700001	3810000
CP000038_38	3800001	3910000

```
CP000038_39          3900000    4010000
CP000038_40          4000000    4110000
CP000038_41          4100000    4210000
CP000038_42          4200000    4310000
CP000038_43          4300000    4410000
CP000038_44          4400000    4510000
CP000038_45          4500000    4610000
CP000038_46          4600000    4710000
CP000038_47          4700000    4810000
CP000038_48          4800000    4825265
Continuation [1] of 49) of CP000038 from base 1000001 (CP000038 Shigella sonnei Ss046,

Alignment Scores:
Pred. No.:
Score:                 1.42e-193
Percent Similarity:    2182.00         Length:        110000
Best Local Similarity: 98.1%           Matches:       424
Query Match:           98.1%           Conservative:  0
DB:                    96.6%           Mismatches:    8
                       15              Indels:        0
                                       Gaps:          0
```

US-10-601-319-10 (1-432) x CP000038_10 (1-110000)

```
Qy     1 MetLysAlaIleLeuIleLeuProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla    20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 60260 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGAATCAAATCTGCA                 60319

Qy    21 PheAlaGlnSerGluProGluLeuHisTyrLeuGluSerValValIleValSerArgHisGly    40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 60320 TTCGGTCAGAGTGAGCCGGAGCTGAAAGTGTGGTAGTGTCAGTCGTCATGGT             60379

Qy    41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro      60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 60380 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA     60439

Qy    61 ThrTrpProValLysGluGlyLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu         80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 60440 ACCTGGCCGGTAAAACTGGGTTGGGCTGACACGCCGGGTGGTGAGCTAATCGCCTATCTC    60499

Qy    81 GlyHisTyrThrProArgGlnArgGlnValValAlaAspGlyLeuLeuProLysCysCysPro   100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 60500 GGACATTACCAACGCCAGGCGGTTGGTAGCCGACGGATTGCTGCCGAAAAGGCTGCCG       60559

Qy   101 GlnSerGlyGlnValValAlaIleLeuIleAlaAspValAspArgThrArgLysThrGlyGlu   120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 60560 CAGTCTGGTCAGGTCGGGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA    60619

Qy   121 AlaPheAlaAlaGlyLeuAlaProAspCysValAlaIleThrValHisThrGlnAlaAspThr   140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 60620 GCCTTCGCCGCCGGGCTGGCACCTGACTGTCAATAACCGTACATACCGGAGGCTGATACG    60679

Qy   141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla     160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 60680 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG    60739

Qy   161 AsnValThrAspAlaIleLeuThrArgAlaAsgPheThrGlyHis                    180
         ||||||||||||||||||||||||||||||||||||||||||||
Db 60740 AACGTGACTGACGCGATCCTCAGCCAGGGGTCAATGCTACTTTTACCGGGCAT          60799

Qy   181 TyrGlnThrAlaAspPheArgGluLeuGlnArgGlnValValLeuAsnPheProGlnSerLeuCys   200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 60800 CGGCAAACGGCGTTCGGACTGGAACGGGTGCTTAATTTTCCGCAAATCAAACTGTGC     60859

Qy   201 LeuLysArgGlyLysGlyAspLeuSerCysSerLeuThrArgGlnAlaLeuProSerGluGlu   220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 60860 CTTAAACGTGGGAAAGGCGATGACGCGTGTTCATTAACGCACGGCCAATACCATCGGAACTC    60919

Qy   221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuLeuAlaSerMetLeuThr   240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 60920 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTTAAGCCTCGCATCAATGCTGACG    60979

Qy   241 GlyIlePheLeuLeuGlnIleAlaGlnGlyMetProGluProGlyThrTrpGlyArgIleThr   260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 60980 GAGATANTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCCGGGACGTGGGGAAGGATCACC   61039
```
```
Qy   261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln      280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61040 GAATTCACACAGTGGAACACCTTGCTAAGTTTGCATAAGCGGCAATTTTATTTGCTACAA    61099

Qy   281 ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuGluLysIleLysThrAla   300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61100 CGCACGCCGGAGGTTGCCGCACGCTCGCGCGCCACCCCGTTATTAGATTTGAAGACAGCG    61159

Qy   301 LeuThrProHisPheProProGlnLysGlnAlaThrGlyValThrLeuProThrSerValLeu   320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61160 TTGACGCCCCATCCACCGCAAAAACAGGCGACGGGTGTGACAATTACCGACTTCAGTGTG    61219

Qy   321 PheIleAlaGlyHisAspThrAlaAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp   340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61220 TTTATCGCCGGACACGATGCTAATCTGGCAAATCTCGGCGGTGCCACTGGAGCTCAACTGG    61279

Qy   341 ThrLeuProGlyProGlnProAspProAsnThrProProGlyGluValGlyLeuValPheGlu   360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61280 ACGGCTTCCCGGTCCAGCCCGGATAACACCCCGCCGGGAGTCGTGGGACTGGTGTTTGAA    61339

Qy   361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerValPheGlnThrLeuGln          380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61340 CGTCGGCGTCCAAGCGACATAACAGCCAGTGGATTTCGCTGCGTTCCAGACTTTACAG    61399

Qy   381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValGlyLeuLeuThr   400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61400 CAGATGCGTGATAAAAACGCCGCTGTCATTAAATACGCGCCGCGGAGGTGAAACTGACC    61459

Qy   401 LeuAlaGlyCysGluGluValArgAsnAlaGlnGlyMetCysSerLeuLeuAlaGlyPheThrGln   420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61460 CTGGCAGGATGTGAAGAGGCGAAATGCCAGGCCCATGTTGTTCGTTGCTGTTGGCAGGT    61519

Qy   421 IleValAsnGluValAlaArgIleLeuProAlaCysSerLeu                        432
         ||||||||||||||||||||||||||||||||||||||||||
Db 61520 ATCGTGAATGAAGCACACGCGATACCGGCCGTGCCAGTTTG                         61555
```

RESULT 15
```
AR636188
LOCUS        AR636188              1901 bp    DNA        linear    PAT    20-APR-2005
DEFINITION   Sequence 5 from patent US 6855365.
ACCESSION    AR636188
VERSION      AR636188.1    GI:62768871
KEYWORDS     Unknown.
SOURCE       Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE    1  (bases 1 to 1901)
  AUTHORS    Short,J.M., Kretz,K.A., Gray,K.A., Batson,N.R., Garrett,J.B.,
             O'Donoghue,E. and Mathur,E.J.
  TITLE      Recombinant bacterial phytases and uses thereof
  JOURNAL    Patent: US 6855365-A 5 15-FEB-2005;
             Diversa Corporation; San Diego, CA
FEATURES             Location/Qualifiers
     source          1..1901
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN
```

```
Alignment Scores:
Pred. No.:
Score:                 4.65e-196
Percent Similarity:    2181.00         Length:        1901
Best Local Similarity: 98.1%           Matches:       424
Query Match:           98.1%           Conservative:  0
DB:                    96.6%           Mismatches:    8
                       2               Indels:        0
                                       Gaps:          0
```

US-10-601-319-10 (1-432) x AR636188 (1-1901)

```
Qy     1 MetLysAlaIleLeuIleLeuProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla    20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA    247
```

```
Qy    21 PheAlaGlnSerGluProGluLeuLysGluSerValValIleValSerArgHisGly    40
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248 TTCGCTCAGAGTGAGCCGGAGCTGAAGAGTGTGGTGATTGTCAGTCGTCATGGT       307

Qy    41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308 GTCCGTGCTCCAACCAAGGCCACCCAACTGATGCAGGATGTCACCCAGACGCAAGGCCA    367

Qy    61 ThrTrpProValLysGluGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu   80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368 ACCTGGCCGGTAAAACTGGGTTGACACCCGGGGTGGTGAGCTAATCGCCTATCTC       427

Qy    81 GlyHisLeuTyrTrpArgGlyLeuValAlaAspGlyLeuLeuProLysGlyGlyCysPro  100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 GGACATTACCAACCGCCAGGGTCTCGGTAGCCGACCGGATTCCTGGGAAAAAGGGCTGCCG  487

Qy   101 GlnSerGlyLeuValAlaIleLeuIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488 CAGTCTCGTGCTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGGGAA 547

Qy   121 AlaPheAlaAlaGlyLeuAlaProAlaAspCysAlaIleThrHisValHisSerGlnAlaAlaAspThr 140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548 GCCTTCGCCGCCGGCCTGGCACCTGCAGATTGTGCCATCACCCATGTTCACTCGCAGGCGGCAGATACG 607

Qy   141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuLeuAspAsnAla 160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   608 TCCAGTCCGGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG   667

Qy   161 AsnValThrAspAlaIleLeuLeuGluValArgThrGlyLysSerIleLeuAlaAspPheThrGlyHis 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668 AACGTGACTGACGCGATCCTCAGCAGGGTCAGGGGCCAGGAGGGTCAAATTGCTGACTTTACCGGGCAT 727

Qy   181 TyrGlnThrAlaPheGluLeuGluGlnValLeuAsnAsnPheProGlnSerArgAsnLeuCys 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   728 CGGCAAACGGCCGTTTCGGGAACTGGAACAGGTGCTTAATAATTTCCCGCAATCAAACTTGTGC 787

Qy   201 LeuLysArgGlyLysGlnAspGluSerCysSerSerLeuThrValAlaLeuProSerGluLeu 220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   788 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGGTAGCATTACCATGGAACTC     847

Qy   221 LysValSerAlaAlaAsnAspCysValSerValSerArgIleAlaAlaValSerLeuLeuThr 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   848 AAGGTGAGCGCCGCCAACAATGTCTCATTAACCGTGCGGTGCCTCGCAATGCTGACG       907

Qy   241 GluIlePheLeuLeuGluGlnGlnAlaGlnGlyMetProGlyProGlyTrpGlyArgIleThr 260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   908 GAGATATTTCTCCTGCAACAGCAGGCGATGCCCGGACCGGGGTGGGGAAGGAATCACC      967

Qy   261 AspSerHisLeuGlnTrpAsnThrLeuLeuSerLeuHisAlaAsnAlaGlnPheAspLeuLeuGln 280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   968 GATTCACACCAGTGGAACACCCTTGCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA   1027

Qy   281 ArgThrProGlyValAlaIleArgSerArgAlaAlaThrProLeuLeuAspLeuIleLysThrAla 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1028 CGCACGCCAGGGGTTGCCCGAGCCGCGGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG   1087

Qy   301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu   320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1088 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCTG   1147

Qy   321 PheIleAlaGlyLysHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1148 TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTCGAGCTCAACTGG    1207

Qy   341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp   360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1208 ACGCTTCCCGGTCAGCCCGGATAACACCCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG   1267

Qy   361 ArgArgLeuSerAspAsnSerGlnThrTrpIleGlnValSerLeuValPheGlnThrLeuGln  380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1268 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG    1327

Qy   381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValGlyLeuThr   400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1328 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCCCCGGGAGAGGTTGGCCTGACC   1387

Qy   401 LeuAlaGlyTyrCysGluGluArgArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1388 CTGGCAGGATGTGAAGAGCGGCGAAATCGCAGGGGCANGTGTGGTCGTTGGCAGGTTTTTACGGCAA 1447

Qy   421 IleValAsnGluValAlaArgIleProAlaAlaCysSerLeuLeu                  432
         |||||||||||||||||||||||||||||||||||||||
Db  1448 ATCGTGAATGAAGCACGCATACCGGCGCCGTGCAGTTTG                       1483
```

Search completed: June 14, 2006, 13:55:38
Job time : 5313.22 secs

THIS PAGE LEFT BLANK

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

Run on: June 14, 2006, 10:38:44 ; Search time 552.385 Seconds
                     (without alignments)
                     8179.104 Million cell updates/sec

Title: US-10-601-319-10
Perfect score: 2258
Sequence: 1 MKAILIPFLSLLIPLTPQSA.........CSLAGPTQIVNEARIPACSL 432

Scoring table: BLOSUM62
                     Xgapop 10.0 ,   Xgapext   0.5
                     Ygapop 10.0 ,   Ygapext   0.5
                     Fgapop   6.0 ,   Fgapext   7.0
                     Delop   6.0 ,   Delext   7.0

Searched: 5244920 seqs, 348612423 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
 N Geneseq 8.*
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abs/ABSWEB_spool/US10601319/runat_12062006_180047_27417/app_query.fasta_1
-DB=N Geneseq -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss06h
-USER=US10601319 @CGN 1 1.749 @runat_12062006_180047_27417 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Databases :

     1: geneseqm1980s:*
     2: geneseqm1990s:*
     3: geneseqm2000s:*
     4: geneseqm2001as:*
     5: geneseqm2001bs:*
     6: geneseqm2002as:*
     7: geneseqm2002bs:*
     8: geneseqm2003as:*
     9: geneseqm2003bs:*
    10: geneseqm2003cs:*
    11: geneseqm2003ds:*
    12: geneseqm2004as:*
    13: geneseqm2004bs:*
    14: geneseqm2005s:*
    15: geneseqm2006s:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2258	100.0	1296	14	AED50817	Aed50817 Escherich
2	2258	100.0	1308	10	ADC87742	Adc87742 DNA encod
3	2258	100.0	1308	12	AD050299	Ad050299 Escherich

ALIGNMENTS

RESULT 1
AED50817

ID   AED50817 standard; DNA; 1296 BP.
XX
AC   AED50817;
XX
DT   29-DEC-2005 (first entry)
XX
DE   Escherichia coli phytase appA mutant DNA SEQ ID NO: 1 #1.
XX
KW   Transgenic animal; gene therapy; feedstuff; food; thermostable; mutant;
KW   ds; coding sequence; 3-phytase; recombinant DNA; appA gene.
XX
OS   Escherichia coli; strain K12.
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1296
FT   /*tag= a
FT   /EC_number= "3.1.3.8"
FT   /partial
FT   /product= "Escherichia coli phytase mutant protein"
FT   /note= "No stop codon"
XX
PN   US2005246780-A1.
XX
PD   03-NOV-2005.

4	2258	100.0				
5		100.0				
6		97.2	1901	12		
7	2190	97.0	1901	12	AED50295	
8	2190	97.0	1901	15	AEE75416	
9	2182	96.6	1299	8	ABK12514	
10	2182	96.6	1299	12	ACA19297	
11	2182	96.6	1299	8	ABK12514	

```
XX
PF   11-FEB-2005; 2005US-00056354.
XX
PR   13-AUG-1997; 97US-00910798.
PR   01-MAR-1999; 99US-00259214.
PR   13-APR-1999; 99US-00291931.
PR   25-MAY-1999; 99US-00315528.
PR   25-MAY-2000; 2000US-00580515.
PR   24-MAY-2001; 2001US-00866379.
PR   24-MAY-2002; 2002US-00156660.
XX
PA   (SHOR/) SHORT J M.
PA   (KRET/) KRETZ K.
PA   (GRAY/) GRAY K A.
PA   (BART/) BARTON N R.
PA   (GARR/) GARRETT J B.
PA   (ODON/) ODONOGHUE E.
PA   (MATH/) MATHUR E J.
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, Odonoghue E;
PI   Mathur EJ;
XX
DR   WPI; 2005-733851/75.
DR   P-PSDB; AED50818.
XX
PT   New isolated or recombinant nucleic acid encoding phytase enzymes, useful
PT   as foodstuff, and for oil degumming, producing an animal feed, delivering
PT   a phytase enzyme supplement to an animal.
XX
PS   Claim 1; SEQ ID NO 1; 104pp; English.
XX
CC   The present invention relates to a recombinant phytase (EC 3.1.2.8)
CC   protein and its encoding also nucleic acid molecule. Phytases of the
CC   invention is a thermostable protein. They are useful in foodstuffs, for
CC   oil degumming, producing an animal feed, delivering a phytase enzyme
CC   supplement to an animal, increasing the resistance of a phytase
CC   polypeptide to enzymatic inactivation in a digestive system of an animal
CC   and in processing of corn and sorghum kernels. The invention is also
CC   useful in gene therapy and in production of transgenic animals. The
CC   present sequence is a Escherichia coli phytase appA mutant DNA. Note: The
CC   present sequence is the SEQ ID NO: 1 which is given in the sequence
CC   listing. This sequence differs from the SEQ ID NO: 1 shown on page 19 in
CC   disclosure of the specification (see AED50824).
XX
SQ   Sequence 1296 BP; 311 A; 342 C; 356 G; 287 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.43e-202
Score:                  2258.00
Percent Similarity:     100.0%
Best Local Similarity:  100.0%
Query Match:            14

Length:        1296
Matches:       432
Conservative:  0
Mismatches:    0
Indels:        0
Gaps:          0
```

```
US-10-601-319-10 (1-432) x AED50817 (1-1296)                RESULT 2
                                                            ADC87742

Qy    1 MetLysAlaIleLeuIlePheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAGCGATCTTAATCCCATTTTAATCTCTTCGATTCCCTTAACCCCGGAATCTGCA 60

Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCCGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCCATGGCCA 180

Qy   61 ThrTrpProValLysGlyGlyGluLeuThrProGlnGlyGlyValLeuIleAlaTyrLeu 80
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ACCTGGCCGGTAAAACTGGGTGAGCTGACACCGCGGGTGGGTGAGCTCAACCTGTGTG 240

Qy   81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GGACATTACTGGCGGCTCAGCGGTTGGTGGCCGGACGGATTGCCCGGATTGGCTGCCCG 300

Qy  101 GlnSerGlyGlnValAlaAlaIleIleLeuAlaAspGluArgThrArgHisThrGluAsn 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGTCTGGTCAGGTCGCCGCGATTATTGCTGACGAGCGTACCCGTCATACAGAGAAT 360

Qy  121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GCCTTCGCCGCCGGGCTGGCACCCGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420

Qy  141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TCCAGTCCCGATCCGTTATTTAATCCCCTAAAAACTGGCGTTTGCCAACTGGATAACGCG 480

Qy  161 AsnValThrAspAlaIleLeuGluAlaGlyGlySerIleAlaAspPheThrGlyHis 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 AACGTGACTGACGCGATCCTCGAGGCAGGGGGGTCAATTGCTGACTTTACGGGGCAT 540

Qy  181 TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnProSerAsnLeuCys 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 TATCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTCCCGAGCAATCAAACTTGTG 600

Qy  201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CTTAAACGTGAGAAACAGGACGAGAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC 660

Qy  221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 AAGGTGAGCGCCGATTGTGTCTCATTAACCGGTGCGGTAAGCCTTGCATCAAATGCTGACG 720

Qy  241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GAGATATTTCTCCTGCAACAGGCACAGGGAATGCCGGAGCCGGGGTGGGGGAGGATCACC 780

Qy  261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GATTCACACCAGTGGAACACCCTTCTAAGTTTGCATAACGCGCAATTTGATTTGCTACAA 840

Qy  281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 900

Qy  301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGCGTGACATTACCGACTTCAGTGCTG 960

Qy  321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 TTTATCGCGGGACACGATACTAATCTGGCAAATCTCGGCGGTGCACTGGAGCTCAACTGG 1020

Qy  341 ThrLeuProGlyGlnProGluAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 ACGCTCCCGGGTCAGCCGGATAACACGCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1080

Qy  361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 CGTCGGCTGAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTGTTCCAGACTTTACAG 1140

Qy  381 GlnMetArgAspGlyThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 CAGATGCGTGATGGAACGCCGCTGTCATTAAATACGCCGCCGGGAGAGGTGAAACTGACC 1200

Qy  401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAAGGCATGTGTTCGTTGGCAGGGTTTTACGCAA 1260

Qy  421 IleValAlaAsnGluIleAlaArgIleProAlaCysSerLeu 432
        ||||||||||||||||||||||||||||||||||||||||||
Db 1261 ATCGTGAATGAAGAGCACCCATACCGGCGTGCTCAGTTTG 1296
```

```
ID   ADC87742 standard; DNA; 1308 BP.
XX
AC   ADC87742;
XX
DT   01-JAN-2004 (first entry)
XX
DE   DNA encoding the modified Escherichia coli phytase.
XX
KW   Phytase; food supplement; enzyme delivery matrix; soybean meal;
KW   thermotolerance; thermostability; kernel; phytate;
KW   myo-inositol-hexaphosphate; inositol; inorganic phosphate;
KW   thermotolerant; feed value; digestion; gene; ds.
XX
OS   Synthetic.
OS   Escherichia coli.
XX
PN   US2003103958-A1.
XX
PD   05-JUN-2003.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1308
FT                   /*tag= a
FT                   /product= "Modified phytase"
XX
PF   24-MAY-2002; 2002US-00156660.
XX
PR   13-AUG-1997;  97US-00910798.
PR   01-MAR-1999;  99US-00259214.
PR   13-APR-1999;  99US-00291931.
PR   25-MAY-1999;  99US-00318828.
PR   25-MAY-2000;  2000US-00580515.
PR   24-MAY-2001;  2001US-00866379.
XX
PS   Claim 9; SEQ ID NO 1; 113pp; English.
XX
PT   New nucleic acid encoding a polypeptide having phytase activity, useful
PT   in improving the feeding value of phytate rich ingredients or as an aid
PT   in phytate digestion.
XX
PA   (DIVE-) DIVERSA CORP.
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI   Mathur EJ;
XX
DR   WPI; 2003-787039/74.
DR   P-PSDB; ADC87743.
XX
CC   The invention discloses a new isolated or recombinant nucleic acid which
CC   encodes a polypeptide having a phytase activity. Also claimed is a
CC   nucleic acid probe, an amplification primer sequence pair, an expression
CC   cassette comprising the nucleic acid, a vector comprising the nucleic
CC   acid, a transgenic non-human animal or plant, or its seed, comprising the
CC   nucleic acid, an antisense oligonucleotide, inhibiting the translation of
CC   a phytase message in a cell, a heterodimer comprising the polypeptide and
CC   a second domain, an array comprising immobilised polypeptide or nucleic
CC   acid, a hybridoma comprising an antibody that specifically binds to the
CC   polypeptide, a food supplement for an animal, an edible enzyme delivery
CC   matrix, an edible pellet comprising a granule edible carrier and the
CC   polypeptide, a .feed composition, a soybean meal, isolating or identifying
CC   the polypeptide, making an anti-phytase antibody, producing a recombinant
CC   polypeptide, determining whether a compound binds to the polypeptide,
CC   identifying a modulator, whole cell engineering of new or modified
CC   phenotypes by using real-time metabolic flux analysis, increasing
CC   thermotolerance or thermostability of the phytase polypeptide, increasing
CC   the resistance of the phytase polypeptide to enzymatic inactivation in a
CC   digestive system and processing of corn and sorghum kernels. The phytase
CC   activity comprises catalysis of phytate (myo-inositol-hexaphosphate) to
CC   inositol and inorganic phosphate and is thermotolerant. The nucleic acid
CC   is useful in improving the feeding value of phytate rich ingredients or
CC   as an aid in phytate digestion.The sequence presented is the DNA
CC   encoding the modified Escherichia coli phytase.
XX
SQ   Sequence 1308 BP; 315 A; 345 C; 357 G; 291 T; 0 U; 0 Other;
```

Alignment Scores:

Pred. No.:		
Score:	3.47e-202	Length: 1308
Percent Similarity:	2258.00	Matches: 432
Best Local Similarity:	100.0%	Conservative: 0
Query Match:	100.0%	Mismatches: 0
DB:	100.0%	Indels: 0
	10	Gaps: 0

```
US-10-601-319-10 (1-432) x ADC87742 (1-1308)

Qy    1  MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla   20
Db    1  ATGAAAGCCATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA   60

Qy   21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly   40
Db   61  TTCGCTCAGAGTGAGCCCGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT  120

Qy   41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro   60
Db  121  GTGCGTGCTCCAACCAAGGCCACCCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA  180

Qy   61  ThrTrpProValLysGluLeuGlyLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu   80
Db  181  ACCTGGCCGGTAAAACTGGGTGACCTGCACCCGCGTGGTGGAGCTGATTGCCTATCTC   240

Qy   81  GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro  100
Db  241  GGACATTACTGGCGTCAGCGTCTGGTTGCTGACGGATTGCTGCCTAAATGTGGCTGCCCG  300

Qy  101  GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu  120
Db  301  CAGTCTGGTCAGGTCGCGATTATTGCTGACGTGGACGAGCGTACCCGTAAAACAGGCGAA  360

Qy  121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr  140
Db  361  GCCTTCGCCGCCGGCTGGCACCTGACGTGCACTACACCCAGGCAGATACG  420

Qy  141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla  160
Db  421  TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACCGGGGTTTGCCAACTGGATAACGCG  480

Qy  161  AsnValThrAspAlaIleLeuGluLysAlaAspGlyLeuArgThrArgTyrThrGlyHis  180
Db  481  AACGTGACTGACGCGATCCTCGAGAGGGCCAGGAGGGTCAAATTGCTGACTTTACCGGGCAT  540

Qy  181  TyrGlnThrAlaPheArgGluLeuGluValLeuAsnProPheProGlnSerAsnLeuCys  200
Db  541  TATCAAACGGCGTTTCGGGAACTCGAACGTGCTTAATCCGTTTCCGCAATCAAACTTTGTGC  600

Qy  201  LeuLysArgGlyLysGlnAspAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220
Db  601  CTTAAACGTGAGAAACAGGACGACGAAAGCTGTTCATTAACGCAGGCATTACCATCCGGAACTC  660

Qy  221  LysValSerAlaAspProCysValSerLeuThrGlyValAlaValSerLeuAlaSerMetLeuThr  240
Db  661  AAGGTGAGCGCCGACCCGTGTGTCTCATTAACCGGTGTGGCGGTAAGCCTCGCATCAATGCTGACG  720

Qy  241  GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGlnLeuTyrTrpGlyTyrArgIleThr  260
Db  721  GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGCAGCTGTATTGGGGATACCGGATCACC  780

Qy  261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln  280
Db  781  GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTGATTTGCTACAA  840

Qy  281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla  300
Db  841  CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG  900

Qy  301  LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu  320
```

```
Db   901  TTGACGCCCATCCACGCAAAAACAGGCGTTATGGTGTGTGACATTACCGACTTCAGTGCTG  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyTyrGlyAlaLeuGluLeuAsnTrp  340

Db   961  TTTATCGCCGGACACGATACTAATCTCGCCAATCTCGGCGCGGCGCTGGAGCTCAACTGG  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   341  ThrLeuProGlyGlnProAspAspAsnThrProGlyGlyGluLeuValPheGluArgTrp  360

Db  1021  ACGCTTCCCGGTCAGCCGGATAACAACACGCCCGGCGGTGAACTGGTGTTCGAACGCTGG  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   361  ArgArgLeuSerAspAsnSerGlnIleTPIleGlnValSerLeuLeuValPheGlnHisThrLeuGln  380

Db  1081  CGTCGCCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGCTCGTTTTCCAGACTTTACGG  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProGlyGlyValValTyrLeuThr  400

Db  1141  CAGATGCGTGATAAAACGCCCGTGTCATTAAATACCGCGGGAGGTGTTGTTTACCTGACC  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   401  LeuAlaGlyLysGluGluArgAsnAlaGlnLneGlyMetCysSerSerLeuAlaGlyPheThrGln  420

Db  1201  CTGGCAGGGAAGTGAAGAGGGAAATGGCCCAGCGGGCATGTGTTCGTTCGTTGGCGAGGTT  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   421  IleValAsnGluAlaArgIleeProAlaCysSerLeu  432

Db  1261  ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTG  1296
```

RESULT 3
AD050299

ID   AD050299   standard; DNA; 1308 BP.
XX
AC   AD050299;
XX
DT   29-JUL-2004   (first entry)
XX
DE   Escherichia coli K12 appA phytase mutant 819PH59 DNA.
XX
KW   AppA phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW   phytate; animal feed; fish feed; dough; baking; gene; mutant; ds.
OS   Escherichia coli; K12.
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FH
CDS   1..1296
                /product= "Mutant phytase"
                /partial
                /note= "Does not include stop codon"
FT   mutation   replace(202,T)
                /*tag= a
                /note= "Corresponds to position 389 in the wild type
                sequence"
FT   mutation   replace(203,G)
                /*tag= b
                /note= "Corresponds to position 390 in the wild type
                sequence"
FT   mutation   replace(250,C)
                /*tag= c
                /note= "Corresponds to position 437 in the wild type
                sequence"
FT   mutation   replace(251,A)
                /*tag= d
                /note= "Corresponds to position 438 in the wild type
                sequence"
FT   mutation   replace(252,A)
                /*tag= e
                /note= "Corresponds to position 439 in the wild type
                sequence"
FT   mutation   replace(255,C)
                /*tag= f
                /note= "Corresponds to position 442 in the wild type
                sequence"
FT   mutation   replace(283,G)
                /*tag= h
                /note= "Corresponds to position 470 in the wild type
                sequence"
FT   mutation   replace(285,G)
                /*tag= i
                /note= "Corresponds to position 472 in the wild type
                sequence"
FT   mutation   replace(289,G)
                /*tag= j
                /note= "Corresponds to position 476 in the wild type
                sequence"
FT   mutation   replace(290,A)
                /*tag= k
                /note= "Corresponds to position 477 in the wild type
                sequence"
FT   mutation   replace(291,G)
                /*tag= l
                /note= "Corresponds to position 478 in the wild type
                sequence"
FT   mutation   replace(502,A)
                /*tag= m
                /note= "Corresponds to position 689 in the wild type
                sequence"
FT   mutation   replace(503,G)
                /*tag= n
                /note= "Corresponds to position 690 in the wild type
                sequence"
FT   mutation   replace(504,C)
                /*tag= o
                /note= "Corresponds to position 691 in the wild type
                sequence"
FT   mutation   replace(541,C)
                /*tag= p
                /note= "Corresponds to position 728 in the wild type
                sequence"
FT   mutation   replace(542,G)
                /*tag= q
                /note= "Corresponds to position 729 in the wild type
                sequence"
FT   mutation   replace(543,G)
                /*tag= r
                /note= "Corresponds to position 730 in the wild type
                sequence"
FT   mutation   replace(676,A)
                /*tag= s
                /note= "Corresponds to position 863 in the wild type
                sequence"
FT   mutation   replace(677,A)
                /*tag= t
                /note= "Corresponds to position 864 in the wild type
                sequence"
FT   mutation   replace(829,T)
                /*tag= u
                /note= "Corresponds to position 1016 in the wild type
                sequence"
XX

PA   (SHOR/) SHORT J M.
PA   (KRET/) KRETZ K.
PA   (GRAY/) GRAY K A.

```
PA   (BART/)  BARTON N R.
PA   (GARR/)  GARRETT J B.
PA   (ODON/)  O'DONOGHUE E.
PA   (MATH/)  MATHER E J.
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI   Mather EJ;
XX
DR   WPI; 2004-374952/35.
DR   P-PSDB; ADO50300.
XX
PT   Producing phytase; involves providing nucleic acid derived from bacteria
PT   encoding polypeptide having phytase activity, and expressing nucleic acid
PT   in yeast.
XX
PS   Claim 3; SEQ ID NO 9; 74pp; English.
XX
CC   The invention relates to a method for producing phytase that involves
CC   providing a nucleic acid encoding phytase derived from a bacteria, and
CC   expressing the nucleic acid in a yeast under conditions that allow
CC   expression of the enzyme in the yeast. The invention also relates to
CC   modified phytase enzyme which has improved thermal tolerance and protease
CC   stability at low pH. The phytase enzyme can be used in foodstuffs to
CC   improve the feeding value of phytate rich ingredients, and in diet of
CC   numerous animals including mammals, fowls and fishes, commercially
CC   significant mammals such as pigs, goats, laboratory rodents, commercially
CC   significant avian species such as chicken, ducks, doves, parrot, etc.,
CC   commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC   etc. in dough making and baking. In dietary aids for animals. The method
CC   provides easy manufacture of the active ingredient loaded biocompatible
CC   composition, higher yields and loading efficiency. The phytase
CC   incorporated in to the dietary aid is safe for animals. The present
CC   sequence is Escherichia coli K12 appA phytase mutant DNA.
XX
SQ   Sequence 1308 BP; 315 A; 345 C; 357 G; 291 T; 0 U; 0 Other;
```

US-10-601-319-10 (1-432) x ADO50299 (1-1308)

Alignment Scores:

```
Pred. No.:
Score:                  3.47e-202        Length:        1308
Score:                  2258.00          Matches:       432
Percent Similarity:     100.0%           Conservative:  0
Best Local Similarity:  100.0%           Mismatches:    0
Query Match:            100.0%           Indels:        0
DB:                     12               Gaps:          0
```

```
Qy         1 MetLysAlaIleIleLeuIlePheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTTGAATCCCGTTAACCCGCAATCTGCA 60

Qy        21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleSerArgHisGly 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGGTGATTGTGTCAGTCGTCATCGT 120

Qy        41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GTGCGTGCTCCAACCAAGGCACGACAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy        61 ThrTrpProValLysGlnLeuThrProGlnGlyGlyGluLeuIleAlaTyrLeu 80
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 ACCTGGCCGGTGAAACAGCTGACCCCGCAGGGTGGCGAACTGATTGCGTATCTC 240

Qy        81 GlyHisTyrTrpArgGlnArgGlyLeuValAlaAspGlyLeuLeuProLysCysGlyPro 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 GGACATTACTGGCGTCAGCGGGGTCTGGTAGCCGACGGATTGCTGCCTAAATGTGGCCCG 300

Qy       101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 CAGTCTGGTCAGGTCGCCATTATTGCTGACGTGGATGAAAGAACTCGTAAAACAGGGGAA 360

Qy       121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 GCCTTCGCCGCGCGGCTGGCTCCGGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420
```

(continued in second column)

```
Qy       141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 TCCAGTCCCGGATCCCGTTATTTAATCTCTAAAAACTGGGCTGTTTGGCAACTGGGAGATA 480

Qy       161 AsnValThrAspAlaIleLeuGluAlaArgAlaGlyGlyProGlyValCysGlnHis 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 AACGTGACTGACGCGATCCTTGAGGGGCGAGGAGGGCGGGTTTGCCAACTGGGGCAT 540

Qy       181 TyrGlnThrAlaPheArgGluLeuGluGlyLeuLysSerIleAlaAsnProSerGluLeu 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 TATCAAACCGCGTTTCGCGAACTGGAACTGGGGTTAAATTTCCGCAATCAAACCTTGTGC 600

Qy       201 LeuLysArgGlyLysGlnAspGlyGluSerCysSerLeuThrGlnAlaLeuProSerGlu 220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 CTTAAACGTGGAAAGCAGGACGGAGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAA 660

Qy       221 LysValSerAlaAspProLysCysValSerLeuThrGlyAlaValSerLeuAlaSerMet 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 AAGGTGAGCGCCGACCTGTGTGTCTCATTAACCGGTGGGGTAAGCCTCGCATCAATGCT 720

Qy       241 GluIlePheLeuLeuGlnGlnValAlaGlnIleGlyMetProGlyGlyTrpGlyArgIle 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 GAGATATTTCTCCTGCAACAAGCACCAGGGAATGCCGGAGGGTGGGGAAGGATCACC 780

Qy       261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 GATTCACCACGTGGAACACCTTGCTAAGTTTGCATAACGCAATTTGATTTGCTACAA 840

Qy       281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 900

Qy       301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCAACTTCAGTGCTG 960

Qy       321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 TTTATCGCCGGGACATGATACAAGCCAGTGGATTCCGGTTCCAGAGCTTTACAG 1140

Qy       341 ThrLeuProGlyGlnProLeuAsnAlaSerProGlyGlyGluLeuValPheGluArgTrp 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 CAGATGCGTGGATAAAAACGGCCGCGTCGTCATTAAAATACGGACCCGAGAGGTGAAACT 1200

Qy       361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 CGTCGGCGTAAGCCGTAAAGCCCAGTGATTCAGGTTTCCCTGGTCTTCCAGGACTTTACAG 1140

Qy       381 GlnMetArgAspGlyThrProLeuSerLeuAsnThrProProGlyGluLeuValLeuThr 400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 CGTCGGCGTAAGCCGTAAAGCCCAGTGATTCAGGTTTCCCTGGTCTTCCAGGACTTTACAG 1140

Qy       401 LeuAlaGlyCysGluGluArgAsnAlaGlnIleGlyMetCysSerLeuAlaGlyPheThrGln 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 CTGGCAGGATGTGGAAGAGCGAAATGCCAGGGCATGTGTTCGTTAGGCAGGTTTTACGCAA 1260

Qy       421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
             ||||||||||||||||||||||||||||||||||||
Db      1261 ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTG 1296
```

```
RESULT 4
AEB75420
ID   AEB75420 standard; DNA; 1308 BP.
XX
AC   AEB75420;
XX
DT   23-FEB-2006  (first entry)
XX
DE   E. coli B modified appA phytase 819PH59 DNA sequence SEQ ID NO: 9.
XX
KW   ds; coding sequence; mutant; thermostable; phytase;
KW   protein stabilization; pharmaceutical; fertiliser; cramp; cell culture;
```

```
KW   osteoporosis; osteopathic; food; appA.
XX
OS   Escherichia coli K12.
OS   Synthetic.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..1296
FT                  /*tag= a
FT                  /product= "modified phytase appA"
XX
PN   US2005281792-A1.
XX
PD   22-DEC-2005.
XX
PF   01-SEP-2004;  2004US-00933115.
XX
PR   13-AUG-1997;  97US-00910798.
PR   01-MAR-1999;  99US-0025921 4.
PR   13-APR-1999;  99US-00291931.
PR   25-MAY-1999;  99US-00318528.
PR   25-MAY-2000;  200US-00580515.
PR   24-MAY-2001;  2001US-0086379.
XX
PA   (SHOR/)  SHORT J M.
PA   (KRET/)  KRETZ K A.
PA   (GRAY/)  GRAY K A.
PA   (BART/)  BARTON N R.
PA   (GARR/)  GARRETT J B.
PA   (ODON/)  ODONOGHUE E.
PA   (BAUM/)  BAUM W.
PA   (ROBE/)  ROBERTSON D E.
PA   (ZORN/)  ZORNER P.
XX
PI   Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, Odonoghue E;
PI   Baum W, Robertson DE, Zorner P;
XX
DR   WPI; 2006-055022/06.
DR   P-PSDB; AEE75421.
XX
PT   Formulation useful as dietary supplement for treating, preventing or
PT   reversing osteoporosis and bone loss, and preventing muscle cramps,
PT   comprises one or more polypeptides having phytase activity.
XX
PS   Example 2; SEQ ID NO 9; 82pp; English.
```

The present sequence is that of a polynucleotide encoding a novel
modified phytase derived from the Escherichia coli appA protein, with
improved thermal tolerance and protease stability compared to the wild-
type. The present invention relates to a novel formulation, comprising at
least one polypeptide having phytase activity, where the polypeptide is a
modified sequence (AEE75421) derived from the Escherichia coli K-12 appA
phytase (AEE75419). The modification of the enzyme, comprising a series
of mutations, improves the thermal tolerance and protease stability of
the protein. The specification also claims a pharmaceutical composition;
a kit; an immobilized phytase; a fertilizer or soil additive; a liquid
supplement for preventing muscle cramps; a hydrating agent; a tissue
culture or cell culture media; and a plant food additive, all comprising
the novel phytase, and a method of reducing pollution and increasing
nutrient availability in an environment or environmental sample by
degrading environmental phytic acid. The novel enzyme is a dietary
supplement useful for treating, preventing or reversing osteoporosis or
bone loss, and preventing muscle cramps. The liquid supplement is useful
for preventing muscle cramps. The method is useful for reducing pollution
and increasing nutrient availability in an environment or environmental
sample by degrading environmental phytic acid, where the environment or
environmental sample comprises a soil or a body of water. The immobilized
phytase is useful in foodstuffs for improving the feeding value of
phytate rich ingredients.

```
SQ   Sequence 1308 BP; 315 A; 345 C; 357 G; 291 T; 0 U; 0 Other;
```

Alignment Scores:
Pred. No.:   3.47e-202   Length: 1308

```
Score:                2258.00        Matches:        432
Percent Similarity:   100.0%         Conservative:    0
Best Local Similarity: 100.0%        Mismatches:      0
Query Match:          100.0%         Indels:          0
DB:                                  Gaps:            0
```

US-10-601-319-10 (1-432) x AEE75420 (1-1308)

```
Qy    1 MetLysAlaIleLeuIlePheLeuSerLeuLeuIleProLeuThrProGlnSerAla   20
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAGCGATCTTAATCCCATTTTTAATCTCTTCTGATTCCGTTAACCCGGAATCTGCA   60

Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly   40
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAACCGGAGCTGAAACTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT   120

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro   60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTGCGTGCTCCAACCAAGGCCACGCAAACTGATGCAGGATGTCACCCCAGACGCCATGCCA   180

Qy   61 ThrTrpProValLysLeuGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu   80
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ACCTGGCCGGTAAAACTGGGGGAGCTGACAACACCGCGCGGCGGTGAGCTAATCGCCTATC   240

Qy   81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro   100
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GGACATTACTGGCGTCAGCGTCTGGTAGCCGACGGATTGCTGCCCAAATGTGGCTGCCCG   300

Qy  101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu   120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGTCTGGTCAGGTCGCGATTATTGCTGACGTCGACGAGCGTACCCGTAAAACAGGCGAA   360

Qy  121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisHisThrGlnAlaAsp   140
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GCCTTCGCCGCGGGGCTGGCACCTGACTGTGCAATAACCGTACATCACACCCAGGCAGATACG   420

Qy  141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla   160
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TCCAGTCCGGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG   480

Qy  161 AsnValThrAspAlaIleLeuGluArgAlaGlyGlySerIleAlaAspPheThrGlyHis   180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 AACGTGACTGACGCGATCCTCGAGCGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT   540

Qy  181 TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys   200
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 TATCAAACCGGCGTTTCCGGAACTGGAACGCGGGTGCTTAATTTTCCGCAATCAAACTTGTGC   600

Qy  201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu   220
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC   660

Qy  221 LysValSerAlaAspProCysValValSerLeuThrGlyValAlaValSerLeuAlaSerLeuThrMetLeuThr   240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 AAGGTGAGCGCCCGACACTGTGTCCATTAACCGGTGCGGTAAGCCTCGATCAATGCTGACG   720

Qy  241 GluIlePheLeuLeuGlnIleAlaGlyMetProGluProGlyTrpGlyArgIleThr   260
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGTGGGGACGGATCACC   780

Qy  261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln   280
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTGATTTGCTGCAA   840

Qy  281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla   300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG   900

Qy  301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu   320
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCAACTTCAGTCGTG   960

Qy  321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp   340
```

```
Db    961  TTTTATCGCGGACGACGATACTAATCTGGCAAATCTCGGGCGGGCCACTCGGAGCTCAACTGG 1020
Qy
Db    341  ThrLeuProGlynProLeuAspAsnThrProProGlyGlyLeuValPheGluArgTrp 360
Qy
Db   1021  ACGCTTCCCGGTCAGCCCGGTAACACGCGCCAGGTGGTGAACTGGGTGTTTGAACGCTGG 1080
Qy
Db    361  ArgArgLeuSerAspAsnSerGlnThrIleGlnValSerLeuuValPheGlnThrLeuGln 380
Qy
Db   1081  CGTCGGCTAAGCGATAACAGCCAGACCATCCAGGTGAGTTCAGGTTTCCGTCTTCCAGACTTTACAG 1140
Qy
Db    381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuuThr 400
Qy
Db   1141  CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCCGGCGGGGTGAAACTGACC 1200
Qy
Db    401  LeuuAlaGlyCysGluGlyLeuArgAsnAlaGlnInGlyMetCysSerLeuuAlaGlyPheThrGln 420
Qy
Db   1201  CTGGCAGGATGTGAAGGGCCGCCGCCGGCGCAGGGCATGTGTTCGTTGGCCAGGTTTTACGGCA 1260
Qy
Db    421  IleValAsnGluAlaArgIleProAlaCysSerLeuu 432
Qy
Db   1261  ATCGTGAATGAAGCACGCGGATACCGGGCGTGCAGTTTG 1296
Qy
```

RESULT 5
AEE75422

```
ID   AEE75422 standard; DNA; 1901 BP.
XX
AC   AEE75422;
XX
DT   23-FEB-2006 (first entry)
XX
DE   E. coli B modified appA phytase 819PH59 DNA sequence SEQ ID NO: 9.
XX
KW   ds; gene; mutant; thermostable; phytase; protein stabilization;
KW   pharmaceutical; fertilizer; cramp; cell culture; osteoporosis;
KW   osteopathic; food; appA.
XX
OS   Escherichia coli K12.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FH
FT   CDS             188..1486
FT                   /product= "modified AppA phytase"
FT   mutation        389
FT                   /replace(389,T)
FT                   /note= "Wild type base T replaced by G"
FT                   /*tag= b
FT   mutation        390
FT                   /replace(390,G)
FT                   /note= "Wild type base G replaced by A"
FT                   /*tag= c
FT   unsure          401..403
FT                   /note= "encodes Arg"
FT                   /*tag= d
FT   mutation        437
FT                   /replace(437,C)
FT                   /note= "Wild type base C replaced by T"
FT                   /*tag= e
FT   mutation        438
FT                   /replace(438,A)
FT                   /note= "Wild type base C replaced by T"
FT                   /*tag= f
FT   mutation        439
FT                   /replace(439,A)
FT                   /note= "Wild type base A replaced by G"
FT                   /*tag= g
FT   mutation        470
FT                   /replace(470,G)
FT                   /note= "Wild type base A replaced by G"
FT                   /*tag= h
FT   mutation        472
FT                   /replace(472,G)
FT                   /note= "Wild type base G replaced by C"
FT                   /*tag= i
FT   mutation        476
FT                   /replace(476,A)
FT                   /note= "Wild type base G replaced by T"
FT                   /*tag= j
FT   mutation        477
FT                   /replace(477,A)
FT                   /note= "Wild type base A replaced by T"
FT                   /*tag= k
FT   mutation
FT                   /replace(478,G)
FT                   /note= "Wild type base A replaced by G"
FT                   /*tag= l
FT   mutation
FT                   /replace(689,A)
FT                   /note= "Wild type base G replaced by T"
FT                   /*tag= m
FT   mutation
FT                   /replace(690,G)
FT                   /note= "Wild type base A replaced by G"
FT                   /*tag= n
FT   mutation
FT                   /replace(691,C)
FT                   /note= "Wild type base G replaced by A"
FT                   /*tag= o
FT   mutation
FT                   /replace(728,C)
FT                   /note= "Wild type base G replaced by T"
FT                   /*tag= p
FT   mutation
FT                   /replace(729,G)
FT                   /note= "Wild type base C replaced by G"
FT                   /*tag= q
FT   mutation
FT                   /replace(730,G)
FT                   /note= "Wild type base G replaced by A"
FT                   /*tag= r
FT   mutation
FT                   /replace(863,A)
FT                   /note= "Wild type base G replaced by T"
FT                   /*tag= s
FT   mutation
FT                   /replace(864,A)
FT                   /note= "Wild type base A replaced by G"
FT                   /*tag= t
FT   mutation
FT                   /replace(1016,T)
FT                   /note= "Wild type base T replaced by G"
FT                   /*tag= u
XX
PN   US2005281792-A1.
XX
PD   22-DEC-2005.
XX
PF   01-SEP-2004; 2004US-00933115.
XX
PR   13-AUG-1997;  97US-00910798.
PR   01-MAR-1999;  99US-00259214.
PR   13-APR-1999;  99US-00291931.
PR   25-MAY-1999;  99US-00318528.
PR   25-MAY-2000;  2000US-00580515.
PR   24-MAY-2001;  2001US-00866379.
XX
PA   (SHOR/)  SHORT J M.
PA   (KRET/)  KRETZ K A.
PA   (GRAY/)  GRAY K A.
PA   (BART/)  BARTON N R.
PA   (GARR/)  GARRETT J B.
PA   (ODON/)  ODONOGHUE E.
PA   (BAUM/)  BAUM W.
PA   (ROBE/)  ROBERTSON D E.
PA   (ZORN/)  ZORNER P.
XX
PI   Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, Odonoghue E;
PI   Baum W, Robertson DE, Zorner P;
XX
DR   WPI; 2006-055022/06.
DR   P-PSDB; AEE75421.
XX
PS   Claim 1; Page; 82pp; English.
XX
PT   Formulation useful as dietary supplement for treating, preventing or
PT   reversing osteoporosis and bone loss, and preventing muscle cramps,
PT   comprises one or more polypeptides having phytase activity.
XX
CC   The present sequence is that of a polynucleotide encoding a novel
CC   modified phytase derived from the Escherichia coli appA protein, with
CC   improved thermal tolerance and protease stability compared to the wild-
CC   type. The present invention relates to a novel formulation, comprising at
CC   least one polypeptide having phytase activity, where the polypeptide is a
CC   modified sequence (AEE75421) derived from the Escherichia coli K-12 appA
```

```
CC    phytase (AEE75419). The modification of the enzyme, comprising a series
CC    of mutations, improves the thermal tolerance and protease stability of
CC    the protein. The specification also claims a pharmaceutical composition;
CC    a kit; an immobilized phytase; a fertilizer or soil additive; a liquid
CC    supplement for preventing muscle cramps; a hydrating agent; a tissue
CC    culture or cell culture media; a plant food additive, all comprising
CC    the novel phytase, and a method of reducing pollution and increasing
CC    nutrient availability in an environment or environmental sample by
CC    degrading environmental phytic acid. The novel enzyme is a dietary
CC    supplement useful for treating, preventing or reversing osteoporosis or
CC    bone loss, and preventing muscle cramps. The method is useful
CC    for preventing muscle cramps. The liquid supplement is useful
CC    and increasing nutrient availability in an environment or environmental
CC    sample by degrading environmental phytic acid, where the environment or
CC    environmental sample comprises a soil or a body of water. The immobilized
CC    phytase is useful in foodstuffs for improving the feeding value of
CC    phytate rich ingredients. NOTE: This sequence is not shown in the
CC    specification but was created from the wild-type sequence (AEE75418) and
CC    the information given in claim 1.
XX
SQ    Sequence 1901 BP; 470 A; 497 C; 500 G; 433 T; 0 U; 1 Other;

US-10-601-319-10 (1-432) x AEE75422 (1-1901)

Alignment Scores:
Pred. No.:
Score:                  2258.00
Percent Similarity:     100.0%
Best Local Similarity:  100.0%
Query Match:            100.0%

DB:
Sequence 1901 BP; 470 A; 497 C; 500 G; 433 T; 0 U; 1 Other;
```

```
Qy    5.85e-202    Length:           1901
Db                 Matches:           432
                   Conservative:        0
                   Mismatches:          0
                   Indels:              0
                   Gaps:                0
```

```
Qy      1  MetLysAlaIleLeuIlePheLeuSerLeuLeuIleProLeuThrProGlnSerAla   20
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188  ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTTCGATTCCGTTAACCCCGCAAAATCTGCA  247

Qy     21  PheLeuAlaGlnSerGluProGluLeuLysLeuGluSerValValIleSerArgHisGly  40
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248  TTCGCTTCAGAGTGAGCCGGAGCTGAAACGTGTGGTCATTAGTCAGTCGTCATGGT  307

Qy     41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnGlyGlyTyrProAlaTrpPro  60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308  GTGCGTGCTCCAACAAAGGCCACGCAACTGATGCAGGATGTCACCCCAGCACGCATGGCCA  367

Qy     61  ThrTrpProValLysLeuGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu  80
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    368  ACCTGGCCGGTAAAACTGGGTGAGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC  427

Qy     81  GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysGlyCysCysPro  100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428  GGACATTACTGGCGCCAGCGTCTGGTAGCCGACGGATTGCTGCCTAAATGTGGCTGCCCG  487

Qy    101  GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488  CAGTCTGGTCAGGTCGCCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA  547

Qy    121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr  140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548  GCCTTCGCCGCCGGGCTGGCACCTGACTGCGCAATCACCGTACATACCCAGGCCAGATACG  607

Qy    141  SerSerProAspProLeuPheAsnProLeuLysSerGlyThrCysGlnLeuAspAsnAla  160
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608  TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG  667

Qy    161  AsnValThrAspAlaIleLeuLysArgAlaGlyGlySerIleAlaAspPheThrGlyHis  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    668  AACGTGACTGACGCGATCCTCGAGAGGGCCGGCGGCAGCATTGCTGACTTTACCGGCAT  727

Qy    181  TyrGlnThrAlaPheArgGluLeuGlyAsnLeuSerAsnLeuCys  200
           |||||||||||||||||||||||||||||||||||||||||||
Db    728  TATCAAACGGCGTTTCGCGAACTGGAACTGAGCAATCTGTGC  787

Qy    201  LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    788  CTTAAAGGTGAGAAACAGGACGAGAAAAGCTGTTCATTAACGACGGCATTACCATCGGAACTC  847

Qy    221  LysValSerAlaAspCysValSerLeuThrGlyAlaValSerMetLeuThr  240
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db    848  AAGGTGAGCGCGGACTGTGTCTCATTAACGGACGGTGCGGTAAGCCTTCGGCATCAATGCTGAGG  907

Qy    241  GluIlePheLeuLeuGlnIleAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr  260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908  GAGATATTTTCTCTGCAACAACCACAGGAATGCCCGAACCCGGGTGGGGAAGGATCACC  967

Qy    261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheLeuLeuGln  280
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968  GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTGATTTGCTACAA  1027

Qy    281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028  CGCACGCCCGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTAGATCTAAGAACGCG  1087

Qy    301  LeuThrProHisProProGlnLysGlnLysAlaTyrGlyValThrLeuProThrSerValLeu  320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088  TTGACGCCGCCATCCACCGCAAAAACAGGCGTATGGGTGACATTACCCACTTCAGTGCTG  1147

Qy    321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148  TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG  1207

Qy    341  ThrLeuProGlyGlnProGluAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208  ACGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG  1267

Qy    361  ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln  380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268  CGTCGGCTGAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG  1327

Qy    381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr  400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328  CAGATGCGTGATAAAACGCCGCTTTCATTAAATACGCCGCCAGGAGAGGTGAAACTGACC  1387

Qy    401  LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln  420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388  CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCCTTGGCAGGGTTTTACGCAA  1447

Qy    421  IleValAlaAsnGluAlaArgIleProAlaCysSerSerLeu  432
           |||||||||||||||||||||||||||||||||||||||||
Db   1448  ATCGTGGAATGAAGCACGCCATACCGGCGTGCAGTTTG  1483
```

```
RESULT 6
AD050296
ID     AD050296 standard; DNA; 1901 BP.
XX
AC     AD050296;
XX
DT     29-JUL-2004  (first entry)
XX
DE     Escherichia coli K12 appA phytase mutant DNA #2.
XX
KW     AppA phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW     phytate; animal feed; fish feed; dough; baking; gene; mutant; ds.
XX
OS     Escherichia coli, K12.
OS     Synthetic.
XX
FH     Key            Location/Qualifiers
FH
FT     mutation       replace(437,C)
FT                    /*tag= a
FT     mutation       replace(438,A)
FT                    /*tag= b
FT     mutation       replace(439,A)
FT                    /*tag= c
XX
PN     US2004091968-A1.
XX
```

```
PD   13-MAY-2004.
XX
PF   20-JUN-2003; 2003US-00601319.
XX
PR   13-AUG-1997; 97US-00910798.
PR   01-MAR-1999; 99US-00259214.
PR   13-APR-1999; 99US-00291931.
PR   25-MAY-1999; 99US-00318528.
PR   25-MAY-2000; 2000US-00580515.
PR   24-MAY-2001; 2001US-0086379.
XX
PA   (SHOR/) SHOR J M.
PA   (KRET/) KRETZ K.
PA   (GRAY/) GRAY K A.
PA   (BART/) BARTON N R.
PA   (GARR/) GARRETT J B.
PA   (ODON/) O'DONOGHUE E.
PA   (MATH/) MATHER E J.
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI   Mather EJ;
XX
DR   WPI; 2004-374952/35.
XX
PT   Producing phytase, involves providing nucleic acid derived from bacteria
PT   encoding polypeptide having phytase activity, and expressing nucleic acid
PT   in yeast.
XX
PS   Disclosure; SEQ ID NO 6; 74pp; English.
XX
CC   The invention relates to a method for producing phytase that involves
CC   providing a nucleic acid encoding phytase derived from a bacteria, and
CC   expressing the nucleic acid in a yeast under conditions that allow
CC   expression of the enzyme in the yeast. The invention also relates to a
CC   modified phytase enzyme which has improved thermal tolerance and protease
CC   stability at low pH. The phytase enzyme can be used in foodstuffs to
CC   improve the feeding value of phytate rich ingredients, and in diet of
CC   numerous animals including mammals, fowls and fishes, commercially
CC   significant mammals such as pigs, goats, laboratory rodents, commercially
CC   significant avian species such as chicken, ducks, doves, parrot, etc.,
CC   commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC   etc. in dough making and baking, in dietary aids for animals. The method
CC   provides easy manufacture of the active ingredient loaded biocompatible
CC   composition, higher yields and loading efficiency. The phytase
CC   incorporated in to the dietary aid is safe for animals. The present
CC   sequence is Escherichia coli K12 appA phytase mutant DNA.
XX
SQ   Sequence 1901 BP; 472 A; 498 C; 501 G; 429 T; 0 U; 1 Other;
```

```
Alignment Scores:

Pred. No.:                4.9e-196
Score:                    2195.00
Percent Similarity:       98.4%          Length:          1901
Best Local Similarity:    98.4%          Matches:         425
Query Match:              97.2%          Conservative:    0
DB:                       12             Mismatches:      7
                                         Indels:          0
                                         Gaps:            0

US-10-601-319-10 (1-432) x AD050296 (1-1901)
```

```
Qy     1  MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla   20
          ||| ||| ||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   188  ATGAAAGCGATCTTAATCCCATTTTTAATCCTTCTGATTCCGTTAACCCCGCAAATCTGCA  247

Qy    21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleSerArgHisGly       40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248  TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGTGGTGATATCGTCGTCANGGT           307

Qy    41  ValAlaGlyAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308  GTGGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCCATGGCCA  367

Qy    61  ThrTrpProValLysLeuGlyLeuGlyLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368  ACCTGGCCGGTAAAAACTGGGTTGGCTGACACCGGCGNGGGTGGTGGTGAGCTAATCGGCCRATCTC 427

Qy    81  GlyHisTyrTrpProGlnArgLeuValAlaAlaAspGlyLeuLeuLeuProLysCysGlyCysPro 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428  GGACANTTACTGGCGCCGCGTCTGGTNGCGGCAGATGGTTGCTGGGCGAAAAAGGGCTGCCCG    487

Qy   101  GlnSerGlyIleValAlaIleIleAlaAspAsnAlaArgThrArgLysThrGlyGlu      120
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488  CAGTCTGGTCAGGTCGCGANTATTGCTGACGAGGGTACCCGTAAAACAGGCGAA         547

Qy   121  AlaPheAlaAlaGlyLeuAlaIleLeuGluAlaGlyGlySerIleAlaAspAspThr       140
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548  GCCTTCGCCGCTGGCCTGGCCATCCTCGAGGCAGGCGGCTCAATTGCTGACGATACG       607

Qy   141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla    160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   608  TCCAGTCCGGATCCGTTATTTAATCCTCTAAAAACTGGGTTTGCCAACTGGATAACGCG    667

Qy   161  AsnValThrAspAlaIleLeuGluGluAlaThrProLeuProLeuProLysLeuAsnTyrAsnAla 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668  AACGTGACTGACGCCATCCTCGAGGAGGCAACGCCGCTCCCACTGCCAAAACTGAACTACAAC 727

Qy   181  TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys    200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   728  TATCAGACCGCCTTCCGGGAACTGGAACGGGTGCTCAATTTTCCGCAATCAAACTTGTGC    787

Qy   201  LeuLysArgGluGlyLysSerCysSerLeuThrGlnAlaLeuProSerGluLeu         220
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   788  CTTAAACGTGAGAAAACAGGACGAAAGCGTTCATTAACCGAAGCTTTACCATCCGAAGTC    847

Qy   221  GluIlePheLeuLeuLeuGlnGlnGlnGlyValProIleTrpGlyTyrGlyGlyIleThr   240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   848  GAGATATTTCTCCTGCAACAAGCAAGGGGTGCCGGAACTGGGGGTGGGAAGGAATCACC    907

Qy   241  LysSerHisArgGlnLeuLeuSerLeuLeuHisAsnAlaGlnGlnProLeuLeuGln      260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   908  AAGGTGAGCGGCCGACAATCTCTCATTAACCGGGTCCGGTAAGCTCCGCATCAANTGCTA  967

Qy   261  AspSerHisGlnGlnThrLeuLeuSerLeuLeuHisAsnAlaGlnGlnProLeuLeuGln   280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   968  GATTCACCACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCGCAATTTTATTGCTACAA 1027

Qy   281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuLeuLysLysThrAla      300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1028  CGCACGCCGGAGGTTGCCCGCAGCCCACCCCGTTATTGAATCAAGAGACACG         1087

Qy   301  LeuThrProHisProProGlnIleLysGlnIleAlaTyrGlyValThrLeuProThrSerVal 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1088  TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCTG 1147

Qy   321  PheIleAlaGlyHisAspThrAsnLeuAsnLeuLeuGlyGlyAlaLeuGluLeuAsnTrp   340
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1148  TTTATCGCCGGACACGATACTAATCTGAACAAATCTGGGGGGGCCCTGGAGCTCAACTGG 1207

Qy   341  ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1208  ACGCTTCCCGGTCAGCCGGATAACACCCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1267

Qy   361  ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln   380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1268  CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1327

Qy   381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr   400
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1328  CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCGGGAGAGGTGAAACTGACC 1387

Qy   401  LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1388  CTGGCAGGATGTGAAGAGCGGAATGCCCAGGGCATGTGTTCCTTGGCAGGTTTTACCCA 1447

Qy   421  IleValAsnGluAlaArgIleProAlaCysSerLeu                          432
          ||||||||||||||||||||||||||||||||||||
Db  1448  ATCGTGAATGAAGCACGCATTACCGGCGTGCAGTTTG                        1483
```

```
RESULT 7
ADO50295
ID   ADO50295 standard; DNA; 1901 BP.
XX
AC   ADO50295;
XX
DT   29-JUL-2004 (first entry)
XX
DE   Escherichia coli K12 appA phytase mutant DNA #1.
XX
KW   AppA phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW   phytate; animal feed; fish feed; dough; baking; gene; mutant; ds.
XX
OS   Escherichia coli; K12.
XX
OS   Synthetic.
XX
FH   Key               Location/Qualifiers
FT   mutation          replace(389,T)
FT                     /*tag= a
FT   mutation          replace(390,G)
FT                     /*tag= b
XX
PN   US20040091968-A1.
XX
PD   13-MAY-2004.
XX
PF   20-JUN-2003; 2003US-0601319.
XX
PR   13-AUG-1997; 97US-00910798.
PR   01-MAR-1999; 99US-00252214.
PR   13-APR-1999; 99US-00291931.
PR   25-MAY-1999; 99US-00318528.
PR   25-MAY-2000; 200US-00580515.
PR   24-MAY-2001; 2001US-00866379.
XX
PA   (SHOR/) SHORT J M.
PA   (KRET/) KRETZ K.
PA   (GRAY/) GRAY K A.
PA   (BART/) BARTON N R.
PA   (GARR/) GARRETT J B.
PA   (ODON/) O'DONOGHUE E.
PA   (MATH/) MATHER E J.
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI   Mather EJ;
XX
DR   WPI; 2004-374952/35.
XX
PT   Producing phytase, involves providing nucleic acid derived from bacteria
PT   encoding polypeptide having phytase activity, and expressing nucleic acid
PT   in yeast.
XX
PS   Disclosure; SEQ ID NO 5; 74pp; English.
XX

CC   The invention relates to a method for producing phytase that involves
CC   providing a nucleic acid encoding phytase derived from a bacteria, and
CC   expressing the nucleic acid in a yeast under conditions that allow
CC   expression of the nucleic acid in the yeast. The invention also relates to
CC   modified phytase enzyme which has improved thermal tolerance and protease
CC   stability at low pH. The phytase enzyme can be used in foodstuffs to
CC   improve the feeding value of phytate rich ingredients, and in diet of
CC   numerous animals including mammals such as pigs, goats, laboratory rodents,
CC   significant mammals such as pigs, goats, laboratory rodents, commercially
CC   significant avian species such as chicken, ducks, doves, parrot, etc.,
CC   commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC   etc.; in dough making and baking, in dietary aids for animals. The method
CC   provides easy manufacture of the active ingredient loaded biocompatible
CC   composition, higher yields and loading efficiency. The phytase
CC   incorporated in to the dietary aid is safe for animals. The present
CC   sequence is Escherichia coli K12 appA phytase mutant DNA.
XX
SQ   Sequence 1901 BP; 475 A; 499 C; 499 G; 427 T; 0 U; 1 Other;
```

```
Alignment Scores:
Pred. No.:
Score:                  1.45e-195
Score:                  2190.00
Percent Similarity:     98.4%
Best Local Similarity:  98.4%
Query Match:            97.0%
DB:                     12

Length:        1901
Matches:       425
Mismatches:    7
Conservative:  0
Indels:        0
Gaps:          0

US-10-601-319-10 (1-432) x ADO50295 (1-1901)

Qy   1   MetLysAlaIleIleLeuIleLeuProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla   20
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 ATGAAAGCGATCTTAATCCCATTTTTAATCTCTTCTGATTCCGTTAACCCGTTAACGCCGCA        247

Qy   21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly          40
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248 TTCGCTCAGAGTGAGCCCGAGCTGAAAGTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT          307

Qy   41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro          60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA          367

Qy   61  ThrTrpProValLysLeuGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu          80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368 ACCTGGCCGGTAAAACTCGGTGAGCTGACACCGCGNGGTGGTGAGCTAATCGCGTACTC          427

Qy   81  GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro          100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 GGACATTACCAACCGCCAGCGTCTGGTGGCAGATGGCCTGCCGAAAAAGTGCGGCTGCCCG        487

Qy   101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu          120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGGGAA        547

Qy   121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr          140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548 GCCTTCGCCGCCGGCCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG        607

Qy   141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla          160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   608 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG        667

Qy   161 AsnValThrAspAlaIleLeuGluArgAlaGlyGlySerIleAlaAspPheThrGlyHis          180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668 AACGTGACTGACGCCATCCTCAGCCAGGGCCAGGAGGGTCAATTGCCTGACTTTACGGGGCAT      727

Qy   181 TyrGlnThrAlaPheGluGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys          200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   728 CGGCAAACGGCGTTTGCCGAACTGGAACGGGTGCTTAATTTCCGGCAATCAAACTTGTGC        787

Qy   201 LeuLysArgGlyLysGlnAspAspGluSerCysSerLeuThrGlnAlaLeuProSerGlu          220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   788 CTTAAACGTGGAAAACAGGACGAAAGCTGTTCATTAACCCAGGCATTACCATCGGAACTC        847

Qy   221 LysValSerAlaAspSerCysSerLeuThrGlyValAlaValSerLeuAlaSerMetLeuThr        240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   848 AAGGTGAGCGCCGACAGTTGCTCATTAACCGGTGTGGCGGTAAGCCTCGCCAATCAATGCTTACAG   907

Qy   241 GluLeuPheLeuLeuGluGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr          260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   908 GAGCTATTTCTCCTGGAACAGGCACAGGGAATGCCGGAGCCCGGCTGGGGAAGGATCACC        967

Qy   261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheLeuLeuGln          280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   968 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTTTATTGCTACAA         1027

Qy   281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla          300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028 CGCACGCCAGAAGTTGCCCGCAGCCGCGCCACCCCGTTATTAGATTTGATCAAGACACG          1087

Qy   301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu          320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 8

ID AEB75416 standard; DNA; 1901 BP.
XX
AC AEB75416;
XX
DT 23-FEB-2006 (first entry)
XX
DE E. coli phytase related DNA sequence SEQ ID NO: 5.
XX
KW ds; thermostable; phytase; protein stabilization; pharmaceutical;
KW fertilizer; cramp; cell culture; osteoporosis; osteopathic; food.
XX
OS Escherichia coli.
XX
PN US2005281792-A1.
XX
PD 22-DEC-2005.
XX
PF 01-SEP-2004; 2004US-00933115.
XX
PR 13-AUG-1997; 97US-0910798.
PR 01-MAR-1999; 99US-00252214.
PR 13-APR-1999; 99US-00291931.
PR 25-MAY-1999; 99US-00318528.
PR 25-MAY-2000; 2000US-00580515.
PR 24-MAY-2001; 2001US-00866379.
XX
PA (SHOR/) SHORT J M.
PA (KRET/) KRETZ K A.
PA (GRAY/) GRAY K A.
PA (BART/) BARTON N R.
PA (GARR/) GARRETT J B.
PA (ODON/) ODONOGHUE E.
PA (BAUM/) BAUM W.
PA (ROBE/) ROBERTSON D E.
PA (ZORN/) ZORNER P.
XX
PI Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, Odonoghue E;
PI Baum W, Robertson DE, Zorner P;
XX
DR WPI; 2006-055022/06.
XX
PT Formulation useful as dietary supplement for treating, preventing or
PT reversing osteoporosis and bone loss, and preventing muscle cramps,
PT comprises one or more polypeptides having phytase activity.

PS Disclosure; SEQ ID NO 5; 82pp; English.
XX
CC The present sequence is a DNA sequence related to the production of the
CC novel thermostable / protease resistant Escherichia coli phytase
CC described in the specification. The present invention relates to a novel
CC formulation, comprising at least one polypeptide having phytase activity,
CC where the polypeptide is a modified sequence (AEB75421) derived from the
CC Escherichia coli K-12 appA phytase (AEB75419). The modification of the
CC enzyme, comprising a series of mutations, improves the thermal tolerance
CC and protease stability of the protein. The specification also claims a
CC pharmaceutical composition; a kit; an immobilized phytase; a fertilizer
CC or soil additive; a liquid supplement for preventing muscle cramps; a
CC hydrating agent; a tissue culture or cell culture media; and a plant food
CC additive, all comprising the novel phytase, and a method of reducing
CC environmental pollution and increasing nutrient availability in an
CC environmental sample by degrading environmental phytic acid. The novel
CC enzyme is a dietary supplement useful for treating, preventing or
CC reversing osteoporosis or bone loss, and preventing muscle cramps. The
CC liquid supplement is useful for preventing muscle cramps. The method is
CC useful for reducing pollution and increasing nutrient availability in an
CC environment or environmental sample by degrading environmental phytic
CC acid, where the environment or environmental sample comprises a soil or a
CC body of water. The immobilized phytase is useful in foodstuffs for
CC improving the feeding value of phytate rich ingredients. NOTE: The
CC present sequence is shown in the sequence listing but not discussed
CC elsewhere in the specification.
XX
SQ Sequence 1901 BP; 475 A; 499 C; 499 G; 427 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 1
Score: 2190.00
Percent Similarity: 98.4% Length: 1901
Best Local Similarity: 98.4% Matches: 425
Query Match: 97.0% Conservative: 0
DB: 15 Mismatches: 7
 Indels: 0
 Gaps: 0

1.45e-195

US-10-601-319-10 (1-432) x AEB75416 (1-1901)

Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 188 ATGAAAGCGATCTTAATCCCATTTTTAATCTTCTGATTCCGTTAAACCCGCAAATCTGCA 247

Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleSerArgHisGly 40
Db 248 TTCGCTCAGAGTGAGCCGGAGCTGAAAGTCGAGAGCGTGGTGATTGTCAGTCGTCATGGT 307

Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAlaAlaTrpPro 60
Db 308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAAGATGTCACCCCAGCAGCGTGGCCA 367

Qy 61 ThrTrpProValLysGluGlyGlyLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 368 ACCTGGCCGGTAAAAGAAGGTGGCGGTTTGACCCCGCGNGGTGGTGAGCTAATCGCTTATC 427

Qy 81 GlyHisTyrTrpArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
Db 428 GGACATTACCAACGCCAGGTCGTGGCAGACGGATTGCTGCCGAAAAAGGGCTGCCCG 487

Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgGlyThrGlyGlu 120
Db 488 CAGTCTGGTCAGGTGGCGATTATTGCTGACGTGGACGAGCGTACCCGTGGAACAGGCGAA 547

Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCTTCGCCGCGGGCCTGGCACCTGACTGCGCAATAACCGTACATACCCAGGCAGATACG 607

Qy 141 SerSerProAspProLeuPheAsnProLeuGlyThrGlyValCysGlnLeuAspAsnAla 160
Db 608 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG 667

Qy 161 AsnValThrAspAlaIleLeuGluArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 668

Db 1088 TTGACGCCCATCCACCCGAAAAACAGCGGTATGGTGTGTGTGACATTACCCACTTCAGTGCTG 1147
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 1148 TTTATCGCCGGACACGATACTAATCTGGCAAATCTGGGCGGGGCACTGGAGCTCAACTGG 1207
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1208 ACGCTTCCCGGTCAGCCCGGATAACACGCCGCCAGGTGGTGAACTGGTTTTGAACGGTGG 1267
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 CCTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1327
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
Db 1328 CAGATGCGTGATAAAACCCCGCTGTCATTAAATACGCCCCCGGGAGGGGTGAAACTGACC 1387
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1388 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA 1447
Qy 421 IleValAsnGluValAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTGAATGAAGCACGGCATACCGGGGGTGCAGTTTG 1483

Db   668 AACGTGACTGACGCGATCCTCAGCAGGGCAGGAGGGGTCAATTGCTGACTTTACCGGGCAT 727

Qy   181 TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   728 CGGCAAACGGGCGGTTTCGGCGAACTGGAACGGGTGCTTAATTTTCCGGCAATCAAACTTGTGC 787

Qy   201 LeuLysArgGlyLysGlnAspGlySerCysSerLeuThrGlnAlaIleLeuProSerGluLeu 220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   788 CTTAAACGTGAGAAACAGGACGGAAAGCTGTTCATTAAACGCAGCHATTAACCATCGGAACTC 847

Qy   221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   848 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCCGCATCAATGCTGACG 907

Qy   241 GluIlePheLeuLeuGlnGlnIleAlaGlnIleAlaGlnMetProGlnTyrTrpGlyArgIleThr 260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   908 GAGATATTTCTCCTGCTGCAACAGACACAGGGAATGCCCCAGTACTGGGGACGGATCACC 967

Qy   261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAlaAsnAlaGlnPheAspLeuLeuGln 280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   968 GATTCACACCAGTGGAACACACTTTGCTAAGTTTGCATTAAACGGCCCAATTTATTTGCTACAA 1027

Qy   281 ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028 CGCACCCCAGAGGTTGCCGCCGCCAGCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 1087

Qy   301 LeuThrProHisProProGlnLysGlyIleAlaGlyAlaTyrGlyValThrProThrSerValLeu 320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 TTGACCCCGCCATCCACCGCAAAAAGCAGGGTATGGGGTGACATTACCCACTTCAGTGCTG 1147

Qy   321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuLeuGlyTyrGlyAlaLeuGluLeuAsnTrp 340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148 TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGGGCACTGGAGCTCAACTGG 1207

Qy   341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValProThrSerValLeu 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208 ACGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGTGGTGAACTGGTGCCTACCAGTGTTGAACGCTGG 1267

Qy   361 ArgArgGlnLeuSerAspAsnSerGlnInTpIleGlnValLeuSerLeuValPheGlnInThrLeuGln 380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268 CGTCGGCCTAAGCGGATAACACGCCGGGATTCAGGTTCGGCTCCTCCAGACTTTACAG 1327

Qy   381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluLeuValLysLeuThr 400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328 CAGATGCGTGATAAAAAGCCGTGTCATTAAATACGCCGCCCGAGGAGGTGAAACTGACC 1387

Qy   401 LeuAlaGlyCysGluGlnArgAspAsnAlaGlyInGlyMetCysSerLeuLeuAlaGlyPheThrGln 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388 CTGGCAGGATGTGAAGAGCGAAATGGCGGGCCAGGGCATGTGTTCGTTGCTGGCAGGGTTTTACCGCAA 1447

Qy   421 IleValAlaAsnGluValArgIleProAlaCysSerLeu 432
          |||||||||||||||||||||||||||||||||||||||
Db   1448 ATCGTGBAATGAAGCACACCCATACCGGCGTGCAGTTTG 1483

Alignment Scores:
Pred. No.:
Score:
Percent Similarity: 98.1%
Best Local Similarity: 98.1%
Query Match: 96.6%
DB: 3

Length: 1299
Matches: 424
Conservative: 0
Mismatches: 8
Indels: 0
Gaps: 0

4.8e-195
2182.00

US-10-601-319-10 (1-432) x ABK12514 (1-1299)

Qy   1 MetLysAlaIleLeuIleLeuProPheLeuSerLeuLeuIleLeuProLeuThrProGlnSerArgAla 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAGCGGATCTTAATCCCATTTTATTCTCTTCTGAATTCTCGTTAACCCGCAATCTGCA 60

Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCCGGAGCTGAAAGTGTGGGTGATTGTCAGTCGTCATGGT 120

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTGCGTGCTCCAACCAAGGCACCACGCAAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy   61 ThrTrpProValLysSerGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ACCTGGCCGGTAAAAACTGGGTTGGCTGACACCCGCGGTGGGTGAGCTAATCGCCTATCTC 240

Qy   81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGCTGCCGAAAAAGGGCCTGCCCG 300

Qy   101 GlnSerGlyGlnValAlaIleIleIleAlaAspValAspProAlaLysThrGlyValGlu 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGTCTGGTCAGGTCGCCGATTAATTGCTGACGTCGACGCGGTACCCGTAAAACAGGCGAA 360

Qy   121 AlaPheAlaAlaGlyLeuAlaProAspAspCysSerAlaIleIleThrValHisThrGlnAlaAspThr 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCCTTCGCCGCGCCCGGGGCTGGCACCTGACTGTGCAATAACCGTACCATACCCAGGCAGATACG 420

Qy   141 SerSerProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAlaAsnAla 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTGTGCCAGTTGGATGCGAATGCG 480

Qy   161 AsnValThrAspAlaIleLeuGluArgAlaGlySerIleAlaAlaAspProThrGlyHis 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AACGTGACTGACGCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGCTGACTTTACCGGTCACGCAT 540

Qy   181 TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db   541 CCGCAAACGGGCGTTCCGGAACTGGACACGGGTCGCTTAATTTCCGCAATCAAACTTGTGC 600
Qy       LeuLysArgGlyLysGlyGlnAspGlyGluSerCysSerSerLeuThrGlnAlaLeuProSerGluLeu
Db   201 LeuLysArgGlyLysGlyGlnAspGlyGluSerCysSerSerLeuThrGlnAlaLeuProSerGluLeu 220
Qy   601 CTTAAACGTGAGAAACAGGACGGAAAGCTGTTCATTAGCCAGAGTGATTACCATCGGAACTC 660
Db       LysValSerAlaAspSerCysValSerLeuThrGlyAlaValSerMetLeuThr
Qy   221 LysValSerAlaAspSerCysValSerLeuThrGlyAlaValSerMetLeuThr 240
Db   661 AAGGTGAGCGCGCCGACAAGTCTCATTACCGGTGCGGTAAGCCTCGCCATCAATGCTGACG 720
Qy   241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGlyProGlyTyrTrpGlyArgIleThr 260
Db   721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGACCCGGGGTGGGGGAAGGATCACC 780
Qy   261 AspSerHisGlnThrTrpAsnThrLeuLeuSerLeuHisSerAsnAlaGlnPheAspLeuGln 280
Db   781 GATTCACACCAGTGGACACACCCTGCTAAGTTTGCATAACGCGAATTTTATTTGCTACAA 840
Qy   281 ArgThrProGluValAlaAlaSerArgAlaThrProLeuLeuAspLeuLysThrArgThrAla 300
Db   841 CGCACGCCAGAGGTTGCCGCGTCACGCGCCACCCCGTTATTAGATTTGAAGACAACGACG 900
Qy   301 LeuThrProHisProProGlnLeuGlyGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db   901 TTGACGCCCCATCCACCGCAAAAACAGGGCTATGGGGTGACATTACCAACTTCAGTGCTG 960
Qy   321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuLeuGlyValAlaLeuGluLeuAsnArg 340
Db   961 TTTATCGCCGGACACGATACTAATCTGGCAAATCTGCGGGTGGCGCTCAACTGG 1020
Qy   341 ThrLeuProGlyGlnProAlaAspAsnThrProGlyGlyValGluLeuPheGluArgTrp 360
Db  1021 ACGCTTCCCGGTCAGCCCGGATAACACCCGCGGCGGTGAACTGGCGTGTTGACGGCTGG 1080
Qy   361 ArgArgLeuSerArgSerGlnThrIleGlnValSerLeuValPheGlnThrLeuGln 380
Db  1081 CGTCGGCGCTAAGCCGATAACAGCCAGTGGATTCAGGTTCCGCGGTCTTCCAGACTTTACAG 1140
Qy   381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProGlyProGlyGluValValLysLeuThr 400
Db  1141 CAGATGCGTGATAAAACCCCGCTGTCATTAAATACGCCGGGACCAGGTGAAAACTGACC 1200
Qy   401 LeuAlaGlyCysGluGluLeuArgAsnAlaGlnGlyMetCysSerSerLeuAlaGlyPheThrGln 420
Db  1201 CTGGCAGGATGTGAAGAGGCAAATGCGCAGGGCANGTGGTTCGGTTGGCAGGTTTTACGGCAA 1260
Qy   421 IleValAlaAsnGluAlaArgIleProAlaCysSerSerLeu 432
Db  1261 ATCGTGAATGAAGCACGGCATACCGGCGTGCAGTTTG 1296
```

RESULT 10
ID   ACA19297 standard; DNA; 1299 BP.
XX
AC   ACA19297;
XX
DT   19-JUN-2003 (first entry)
XX
DE   Prokaryotic essential gene #954.
XX
KW   Antisense; ds; prokaryotic essential gene; cell proliferation;
KW   drug design; gene.
XX
OS   Escherichia coli.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI   Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
PT   WPI; 2003-029926/02.
DR   P-PSDB; ABU15427.
XX
PS   Claim 14; SEQ ID NO 7167; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC   prokaryotic essential genes. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1299 BP; 317 A; 345 C; 355 G; 282 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:
Score: 4.8e-195   Length: 1299
                       2182.00   Matches: 424
Percent Similarity: 98.1%   Conservative: 0
Best Local Similarity: 98.1%   Mismatches: 8
Query Match: 96.6%   Indels: 0
DB: 8   Gaps: 0

```
US-10-601-319-10 (1-432) x ACA19297 (1-1299)

Qy      1 MetLysAlaIleIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAGCGATCATTAAATCCCATTTATTTAACTCTCTGATTCCGTTAACCCCGCAATCTGCA 60

Qy     21 PheAlaGlnSerGluProGluLeuTyrLeuSerGlnSerValValIleValSerArgHisGly 40
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCGCTCAGAGTGAGCCGGAGCTGTATCTGAGCCAGAGTGTGGTGATTGTCAGTCGTCATGGT 120

Qy     41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGGCGGCTCCAACCAAGGCCACCCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180
```

```
Qy    61  ThrTrpProValLysGluLeuGlyGluThrProArgGlyGlyGluLeuIleAlaTyrLeu    80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  ACCTGGCCCGTGAAACTGGGGGTTGGCTGACACCGCGGGGTGGTGAGCTAATCGCTTATCTC   240

Qy    81  GlyHisTyrTrpProArgGlnArgLeuValAlaLeuAlaAlaGlyLeuLeuProLysCysCysPro  100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  GGACAATTACCAACGCAGCCAGCGTCTGGTAGCCCACGGATTGCTGGGGGGAAAAGGGCTGCCCG   300

Qy   101  GlnSerGlyGlnValAlaIleIleIleLeuAlaAspValAlaAspGlyArgThrArgLysThrGlyGlu  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  CAGTCTGGTCAGGTCGCCGGATTATTGCTAGCGGTAGTCGACGACGGTACCCGTAAAACAGGCGGAA   360

Qy   121  AlaPheAlaAlaGlyLeuAlaProAspCysSerAlaIleIleThrArgValHisThrGlnAlaAspThr  140
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  GCCCTTCGCCGCCGGGCTGGCACCTGACGTGCCAATAACCGTACATACCCAGGCACGATACG   420

Qy   141  SerSerProAspProLeuPheAsnProLeuLysCysThrGlyValCysGlnLeuAspAsnAla  160
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   421  TCCAGTCCCGATCCGTTATTAATCCTTCTAAAAAACTGGTTTGCCAACTGGATAACGCG   480

Qy   161  AsnValThrAspAlaIleLeuGluArgAlaGlyGlySerIleAlaAspPheThrGlyHis  180
          |||||||||||||||||||||||||||||||||||||||||||||
Db   481  AACGTGACTGACGCGGATCCCTCAGCACGGGCAGGGGTCAATTGCCGACTTTACCGGGCAT   540

Qy   181  TyrGlnThrAlaPheArgGluLeuGlyArgValLeuAsnPheProGlnSerAsnLeuCys  200
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   541  CGGCAAAACGGCGGTTTCGGCGAACTGGGACGCGTGCTTAATTTTCCGCAAATCAAACTTGTGC   600

Qy   201  LeuLysArgGluValGlyAspAlaGluSerLeuTrpThrGlnAlaIleProSerGluLeu  220
          |||||||||||||||||||||||||||||||||||||||||||||
Db   601  CTTAAACGTGAGAAAGGCGATCCTCAGCACGGGCAGGGGTCAATTGCTGACTTTACCGGGCAT   660

Qy   221  LysValSerAlaAspCysValSerLeuThrGlyValAlaValAlaSerLeuAlaSerMetLeuThr  240
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   661  AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGGATGCCTCGGCCATCAATGCTGACG   720

Qy   241  GluIlePheLeuLeuGlnAlaGlyGlyMetProGluProGlyTyrTrpGlyArgIleThr  260
          ||||||||||||||||||||||||||||||||||||||||||||
Db   721  GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGTGGGGAAGGATCACC   780

Qy   261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaLeuAlaGlnPheAspLeuLeuGln  280
          |||||||||||||||||||||||||||||||||||||||||||
Db   781  GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCCCAATTTTATTTGCTACAA   840

Qy   281  ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla  300
          ||||||||||||||||||||||||||||||||||||||||||
Db   841  CGCACGCCAGAGGTTGCCGCACGCTCCCGTGCCACGCCGTTATTAGATTTGATCAAGACAGCG   900

Qy   301  LeuThrProHisProProGlyLysGlnAlaTyrGlyValThrLeuProThrThrSerValLeu  320
          |||||||||||||||||||||||||||||||||||||||||||
Db   901  TTGACGCCCCATCCACCGGGCAAAAACAGGCGTATGGTGTGACATTACCAACCACTTCAGTCCTG   960

Qy   321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340
          ||||||||||||||||||||||||||||||||||||||||||||
Db   961  TTTATCGCCGGACACGATACTAATCTGGCGAATCTCGGCGGCGCACTGGAGCTCAACTGG  1020

Qy   341  ThrLeuProGlyGlnProAspAsnThrProProGlyGluValLeuValPheGluArgTrp  360
          ||||||||||||||||||||||||||||||||||||||||||
Db  1021  ACGCTTCCCGGTCAGCCCGATAACACCCCGCCAGGTGAACGTGGTTGAACGGCTGG  1080

Qy   361  ArgArgLeuSerAspAsnSerGlnIleTrpIleLeuAsnValSerLeuValPheGlnThrLeuGln  380
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTTCAGACTTTACAG  1140

Qy   381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLeuValPheLeuThr  400
          |||||||||||||||||||||||||||||||||||||||||||||
Db  1141  CAGATGCGTGATAAAACGCCGCTGTCATTAAATACCCCGCCCGGAGAGGTGAAACTGACC  1200

Qy   401  LeuAlaGlyCysGluGlyLeuArgAsnAlaGlnGlnGlyMetCysSerLeuAlaGlyPheThrGln  420
          |||||||||||||||||||||||||||||||||||||||||||||
Db  1201  CTGGCAGGGATGTGAAGCGGGAAATGCCCAGGGCCATGTGTCCCTGGCAGGTTTTACGCAA  1260

Qy   421  IleValAsnGluValArgIleProAlaCysSerLeu  432
          |||||||||||||||||||||||||||||||||
Db  1261  ATCGTGAATGAAGCACCGCATACCGGCGGTGCAGTTTG  1296
```

RESULT 11

```
ID   ADL16137 standard; DNA; 1299 BP.
XX
AC   ADL16137;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Escherichia coli K-12 AppA phytase gene.
XX
KW   AppA phytase; modified phytase; K-12; animal feed; feed additive;
KW   phosphorus; phytate; myo-inositol hexakisphosphate; food production;
KW   wet grain milling; dental care product; bone resorption; osteoporosis;
KW   osteopathic; kidney stone; metal removal; waste treatment;
KW   cleaning product; rust removal product; gene; ds.
XX
OS   Escherichia coli; K-12.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..1299
FT                  /*tag= a
FT                  /product= "AppA phytase"
XX
PN   WO2004015084-A2.
XX
PD   19-FEB-2004.
XX
PF   11-AUG-2003; 2003WO-US025058.
XX
PR   12-AUG-2002; 2002US-0403330P.
XX
PA   (GEMV ) GENENCOR INT INC.
XX
PI   Blattmann BO, Darzins A, Davis JM, Encell LP, Morrison TB;
PI   Mrachko GT, Schellenberger V;
XX
DR   WPI; 2004-169498/16.
DR   P-PSDB; ADL16129.
XX
PT   Novel recombinant phytase activity comprising modified phytase activity comprising
PT   modification of amino acid residues in mature Escherichia coli phytase
PT   designated EBC18B2, useful as feed additive.
XX
PS   Claim 19; Fig 21; 107pp; English.
XX
CC   The invention relates to recombinant modified AppA phytases having
CC   modified phytase activity. The AppA phytases are derived from a mature
CC   Escherichia coli AppA phytase designated EBC18B2 (see ADL16120), and
CC   comprise substitutions at least one of residues 26, 43, 46, 54, 73, 113,
CC   126, 184, 228, 384 or 410, or at a residue located within 5 residues of
CC   those listed above. The invention also encompasses polynucleotides
CC   encoding a modified AppA phytase and which further encodes a secretion
CC   signal sequence operable in Bacillus sp.; expression constructs, vectors
CC   and host cells (preferably Bacillus subtilis) comprising the modified
CC   polynucleotide sequences; and a method for the production of the modified
CC   AppA phytases. The invention also discloses a method of producing an
CC   enzyme having phytase activity in Bacillus subtilis; reducing the
CC   level of phosphorus in animal manure; a composition comprising a modified
CC   AppA phytase; and animal feed comprising a modified AppA phytase. The
CC   including seeds, capable of expressing a modified AppA phytase. The
CC   modified AppA phytases are useful as an animal feed additive,
CC   particularly in feeds for non-ruminant livestock such as pigs and poultry
CC   which lack the digestive enzymes for extracting phosphorus from phytate
CC   (myo-inositol hexakisphosphate), a substance found in plant seed which is
CC   an important source of phosphorus. The modified AppA phytases also have
CC   many other agricultural, industrial, medical and nutritional
CC   applications. For example, they can be used in the isolation and
CC   recovery of rare metals to produce lower phosphate homologs of phytate,
```

```
CC   which may be used in toothpaste and other dental care products as well as
CC   potential treatments or preventatives of bone resorption (e.g., in
CC   osteoporosis) and renal calculi (kidney stones); in food production
CC   (e.g., tofu and nutritional supplements); in wet grain milling; in
CC   cleaning, personal care, medical and nutritional products; in rust
CC   removal products; and in the removal of metals and other polycations from
CC   such diverse materials as waste products and carbonated beverages. The
CC   present sequence represents Escherichia coli strain K-12 AppA phytase
CC   gene.
XX
SQ   Sequence 1299 BP; 317 A; 345 C; 355 G; 282 T; 0 U; 0 Other;
```

Alignment Scores:

Pred. No.:		
Score:	4.8e-195	
	2182.00	Length: 1299
Percent Similarity:	98.1%	Matches: 424
Best Local Similarity:	98.1%	Conservative:
Query Match:	96.6%	Mismatches: 8
DB:	12	Indels: 0
		Gaps: 0

```
US-10-601-319-10 (1-432) x AD16137 (1-1299)

Qy    1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ARGAAAGGGATCTTAATCCCATTTTAATCTCTTCTGATTCCGTTAACCCCGAATCTGCA 60

Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleSerArgHisGly 40
        ||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTCATTAGTCGTCATGGT 120

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTGCGTGCTCCAACCAAGGCACAGCTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy   61 ThrTrpProValLysGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ACCTGGCCGGTAAAAGGGGTTGGCTGACACCCCGCGGTGGTGAGCTAATCGCCTATCTC 240

Qy   81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GGACATTACCAACGGCAGCGTCTTGGTAGCCGACGGATTGCTGCCGAAAAAGGCTGCCCG 300

Qy  101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGTCTCGGTCAGGTCGCCGAATTATTGCTGATGTCGACGAGCGCACCCGTAAAAACAGGCGAA 360

Qy  121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420

Qy  141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TCCAGTCCCGATCCGTTATTTAATCCCTTAAAAACTGGCGTTTGCCAACTGGATAACGCG 480

Qy  161 AsnValThrAspAlaIleLeuGluArgAlaGlySerIleAlaAspPheThrGlyHis 180
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  481 AACGTGACTGACGCGATCCTCAGCAGGGCAGGGTCAATTGCTGACTTTACCGGGCAT 540

Qy  181 TyrGlnThrAlaLeuPheArgGluLeuGluSerValLeuAsnPheProGlnSerAsnLeuCys 200
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  541 CGGGCAAACGCGCGGTTTCCGGAACGGGTTCTTAATTTTCCGCAATCAAACTTGTGC 600

Qy  201 LeuLysArgGlyLysGlnAspSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CTTAAACGTGAGAAACAGGACAGCTGTTCATTAACGACAGGCATTACCATCGGAACTC 660

Qy  221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  661 AAGGTGAGCGCCGACATGTCTCATTAACGGTGCGGTTAAGCCTCCGATCAATGCTAACG 720

Qy  241 GluIlePheLeuLeuGlnGlnAlaGlnGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GAGATATTTCTCCTGCAACAGGCACAGCAGGGAATGCCGGAGCCGGGTGGGGAAGGGAATCACC 780

Qy  261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GATTCACACCAGTGGAACACACTTGCTAAGTTGCATAACCGGCAATTTAATTTGCTACAA 840

Qy  281 ArgThrProGluValAlaLeuArgGlySerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  841 CGCACGCCAGAGGTTGCCCTGCGCGGCAGCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 900

Qy  301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TTGACGCCGCCCCATCCACCGCAAAAGCAGGCGTATGGTGTGACATTACCCACTTCAGTGCTG 960

Qy  321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  961 TTTATCGCCGGCCACGATACCAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG 1020

Qy  341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
        ||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 ACGCTTCCCGGTCAGCCCGGATAACACGCCGCCAGGTGGTGAACGGTGTGTTTGAACGCTGG 1080

Qy  361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
        ||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 CGTCGGCTGAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140

Qy  381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
        ||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCCGGAGAGGTGAAACTGACC 1200

Qy  401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
        ||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 CTGGCAGGATGTGAAGAGCGGAATGCGCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA 1260

Qy  421 IleValAsnGluLeuArgIleProAlaCysSerLeu 432
        |||||||||||||||||||||||||||||||
Db 1261 ATCGTGAATGAACGCACGCATACCCGCGTGCAGTTTG 1296
```

```
RESULT 12
AAD25463
ID   AAD25463 standard; DNA; 1901 BP.
XX
AC   AAD25463;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Escherichia coli appA phytase wild type DNA.
XX
KW   Bacterial phytase; K12 appA phytase; protease stability; anabolic;
KW   gastrointestinal; nutritional value; feed treatment process; therapy;
KW   thermal tolerance; growth performance; alcoholic drink; biopulping;
KW   non-alcoholic drink; biobleaching; ds.
XX
OS   Escherichia coli.
XX
FH   Key             Location/Qualifiers
FH
FT   CDS             188..1486
FT                   /tag= a
FT                   /product= "E. coli appA phytase protein"
XX
PN   WO200190333-A2.
XX
PD   29-NOV-2001.
XX
PF   24-MAY-2001; 2001WO-US017118.
XX
PR   25-MAY-2000; 2000US-00580515.
XX
PA   (DIVE-) DIVERSA CORP.
XX
PI   Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, O'donoghue E;
XX
DR   WPI; 2002-083108/11.
DR   P-PSDB; AAB15807.
```

XX
PT New bacterial phytase for e.g. improving the nutritional value of phytate
PT -containing foodstuffs and subsequently improving the growth performance
PT of an organism that consumes it, or in treating animal digestive systems.
XX
PS Claim 42; Fig 7; 170pp; English.
XX
CC The patent discloses recombinant bacterial phytase from Escherichia coli
CC K12 appA phytase. The enzyme has phytase activity and improved thermal
CC tolerance when compared with wild-type phytase. It has improved protease
CC stability at low pH. The recombinant phytase is useful for improving the
CC nutritional value of phytate-containing foodstuffs and subsequently
CC improving the growth performance of an organism that consumes it, in
CC treating animal digestive systems. In feed treatment processes and for in
CC vitro purposes related to research, discovery and development. They are
CC also used for generating recombinant digestive system life forms, for
CC producing or manufacturing alcoholic and non-alcoholic drinks based on
CC the use of moulds, grains and/or plants, in biopulping and bio- bleaching
CC where a reduction in the use of environmentally harmful chemicals that
CC are traditionally used in the pulp and paper industry is desired and in
CC the reduction or possible elimination of the need for mineral
CC supplements, enzymes or therapeutic drugs for animals from the daily feed
CC thus increasing the amount calories and nutrients present in the feed.
CC The present sequence is a DNA encoding E. coli appA phytase wild type
CC protein
XX
SQ Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:
Score: 8.17e-195
Percent Similarity: 2182.00
Best Local Similarity: 98.1%
Query Match: 98.1%
DB: 96.6%
                           6

                           Length: 1901
                           Matches: 1901
                           Conservative: 424
                           Mismatches: 0
                           Indels: 8
                           Gaps: 0

US-10-601-319-10 (1-432) x AAD25463 (1-1901)

Qy   1 MetLysAlaIleLeuLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 ATGAAAGCGATCTTAATCCCATTTTTATTCTCTTCTGATTCCGTTAACCCCCGCAAATCTGCA 247

Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248 TTCGCTCAGAGTGAGCCCGAGCTGAAACGTGGGAAAGTGTCAGTTCGTCATGGT 307

Qy   41 ValArgAlaProThrLysAlaThrHisGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308 GTGCGTGCTCCAACCAAGGCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 367

Qy   61 ThrTrpProValLysGluGlyLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368 ACCTGGCCGGTAAAACTGGGTTGCGTGACACCGCGGGGTGGTGAGCTAATCGCCTATCTC 427

Qy   81 GlyHisTyrTrpProArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGCTGCCGAAAAAGGGCTGCCCG 487

Qy   101 GlnSerGlyInValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488 CAGTCTGGTCAGGTGGCCATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA 547

Qy   121 AlaPheAlaGlyLeuAlaProAspGlyAlaIleThrValHisThrGlnAlaAspThr 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548 GCCTTCGCCGGCGGCGTGGCCACCTGACGTCGCAATAACCGTACATACCCAGGCAGATACG 607

Qy   141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   608 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAACCGGGTGTGCAACTGGATAACGCG 667

Qy   161 AsnValThrArgAlaIleLeuGluAlaArgGlyAlaGlyGlySerIleAlaAspPheThrGlyHis 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668 AACGTGACTCGCGCGATCCTCAGCAGGGCACGGGGGGTCAATGCTGACTTTACGGGCAT 727

Qy   181 TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   728 CGGCAACAGGCGGTTTCGCGAACTGGAACGGGGTGCTTAATTTTCCGCAATCAAAACTTGTGC 787

Qy   201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaAlaLeuProSerGluLeu 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   788 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCAGCATTACCATCGGAACTC 847

Qy   221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   848 AAGGTGAGCGCCGACAAATGTCATTAACCGGTCCGGTAAGCCTGGCTTCGATGCTGACG 907

Qy   241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   908 GAGATATTTCTCCTGCAACAGGCACAGGGAATGCCGGAACCGGGCTGGGGAAGGATCACC 967

Qy   261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   968 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCCCAATTTGATCTACA 1027

Qy   281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028 CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 1087

Qy   301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGGGTGACATTACCCACTCAGTGCTG 1147

Qy   321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148 TTTATCGCCGGACACGATACCAATCTGGCAAATCTCGGCGGTGCACTGGAGCTCAACTGG 1207

Qy   341 ThrLeuProGlyGlnProProGlyGlyGluLeuValPheGluArgTrp 360
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   1208 ACGCTTCCGGGTCAGCCGCCGGGATAACACCGCCCGGAACTGGTTTGAACGCTGG 1267

Qy   361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268 CGTCGGCTGAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1327

Qy   381 GlnMetArgAspProLeuProLeuSerLeuAsnThrProLeuGluGluLeuGlyGluLeuThr 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328 CAGATGCGTGATCAAAACGCCGCTGTCATTAAATACGGCCGCCAGAGGTGAAACTGACC 1387

Qy   401 LeuAlaGlyCysGluGluLeuArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388 CTGGCACGGAATGTGAAGAGCGAAATGGGCGATGTGTTCGTTGGCAGGTTTTACGGCAA 1447

Qy   421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
          ||||||||||||||||||||||||||||||||||||
Db   1448 ATCGTGAATGAAGCAGCCGCCATACCGGCGTGCAGTTTG 1483

RESULT 13
ADA19449
ID   ADA19449 standard; DNA; 1901 BP.
XX
AC   ADA19449;
XX
DT   20-NOV-2003 (first entry)
XX
DE   E. coli K12 appA gene encoding phytase.
XX
KW   Phytase; da; gene; phytate; appA gene; animal feed; inorganic phosphate;
KW   digestion enhancement; transgenic; thermal tolerance; protease stability.
OS   Escherichia coli; strain K12.
XX
FH   Key   Location/Qualifiers
FT   CDS   188..1486
FT   /*tag= a
FT   /product= "Phytase"
XX

```
PN    US2002136754-A1.
XX
PD    26-SEP-2002.
XX
PF    24-MAY-2001; 2001US-00866379.
XX
PR    13-AUG-1997; 97US-00910798.
PR    01-MAR-1999; 99US-00259214.
PR    13-APR-1999; 99US-00291931.
PR    25-MAY-1999; 99US-00318528.
PR    25-MAY-2000; 2000US-00580515.
XX
PA    (SHOR/) SHORT J M.
PA    (KRET/) KRETZ K A.
PA    (GRAY/) GRAY K A.
PA    (BART/) BARTON N R.
PA    (GARR/) GARRETT J B.
PA    (DONO/) O' DONOGHUE E.
PA    (MATH/) MATHUR E J.
XX
PI    Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB;
PI    O' Donoghue E, Mathur EJ;
XX
PT    Isolated Escherichia coli polynucleotide encoding a modified phytase
PT    enzyme, useful in the production of animal feed, for improving the
PT    nutritional value of phytate-containing foodstuff and for enhancing
PT    digestion in humans and animals.
XX
PS    Claim 42; Fig 7; 62pp; English.
XX
CC    The invention relates to an isolated Escherichia coli polynucleotide
CC    encoding a phytase enzyme appearing as ADA19450 and having amino acids
CC    modification from W68R, Q84W, A95P, K97C, S168E, R180Y, N226C or Y277D.
CC    Also included the E. coli appA gene ADA19449 (or an oligonucleotide
CC    derived from it) or its mutant sequence ADA19452, expression vectors,
CC    host cells, a method of improving nutritional value of a phytate-
CC    containing foodstuff by contacting the phytate-containing foodstuff with
CC    a pure phytase enzyme mutated phytase (where the phytase enzyme catalyzes
CC    the liberation of inorganic phosphate from the phytate in the phytate-
CC    containing foodstuff), a method to produce an animal feed containing a
CC    microbial phytase (comprising culturing the plant cell, plant part or
CC    plant under conditions where the nucleotide sequence is expressed and
CC    converting the plant cells, plants or plant into a composition for animal
CC    feed), a feed composition for animals (comprising the plant seeds, plant
CC    cells, plant parts or plants in admixture with a phytate-containing
CC    foodstuff), a method to treat a human or an animal able to benefit from
CC    digestive enhancement by the activity of an exogenous phytase enzymes
CC    comprising administering to the human or animal the plant seed, plant
CC    cells, plant parts or plants of a transgenic plant which is modified to
CC    contain an expression system which expresses a nucleotide sequence
CC    encoding a phytase enzyme, a transgenic non-human organism whose genome
CC    comprising a heterologous nucleic acid sequence encoding a polypeptide
CC    having phytase activity. The phytase enzyme is useful for improving the
CC    nutritional value of phytate-containing foodstuff, in the production of
CC    animal feed and for enhancing digestion in humans and animals. The
CC    invented method improves thermal tolerance and protease stability. It
CC    also improves the feeding value of phytate rich ingredients. The present
CC    sequence represents the E. coli K12 appA gene encoding wild-type phytase.
XX
SQ    Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;
```

```
Alignment Scores:
Pred. No.:              8.17e-195
Score:                  2182.00
Percent Similarity:     98.1%
Best Local Similarity:  98.1%
Query Match:            96.6%

                Length:     Matches:
                1901        1901
                            424
                Conservative: 0
                Mismatches:   8
                Indels:       0
                Gaps:         0

DB:
US-10-601-319-10 (1-432) x ADA19449 (1-1901)
```

DR   WPI; 2003-040002/03.
DR   P-PSDB; ADA19450.

```
Qy      1 MetLysAlaIleIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
        188 ATGAAAGCGATCATTAATCCCATTTTTATCTCTTCTGAATTCCGTTAACCCCGCAAATCTGCA 247

Qy     21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
        248 TTCGCTCAGAGTGAGCCGGAGCTGAAAGCTGGAAAGTGTGGTGATTGTCAGTCGTGTCATGGT 307

Qy     41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
        308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 367

Qy     61 ThrTrpProValLysGluLeuGlyLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
        368 ACCTGGCCGGTAAAAGAACTGGGTCTGACACCGCGTGGCGGTGAGCTAATCGCCTATCTC 427

Qy     81 GlyHisTyrTrpArgGlnArgLeuValAlaAlaAspGlyLeuLeuProLysCysGlyCysPro 100
        428 GGACATTACTGGCGCCAGCGTCTCGTTGCTGCCGACGGATTGCTGCCGAAAAAGGGCTGCCCG 487

Qy    101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
        488 CAGTCTGGTCAGGTCGCGATTATTGCCGATGTCGACGAACCCACGCGTAAAACAGGCGAA 547

Qy    121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
        548 GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 607

Qy    141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
        608 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGGGTTTGCCAACTGGATAACGCG 667

Qy    161 AsnValThrAspAlaIleLeuGluArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
        668 AACGTGACTGACGCGATCCTTGAGCGGGCCGGTGGCTCAATTGCTGACTTTACCGGGCAT 727

Qy    181 TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnProSerGlnAsnAsnLeuCys 200
        728 CGGCAAACCGCGTTTCGCGAACTGGAACGGGTGCTTAATTTCCCAATCAAACTTGTGC 787

Qy    201 LeuLysArgGlyLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
        788 CTTAAACGTGGAAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC 847

Qy    221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
        848 AAGGTGAGCGCCGACAAATGTCTCATTAACCGTGGCGGTAAGCCTCGCATCAAATGCTGACG 907

Qy    241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGlnThrTrpGlyTyrArgIleThr 260
        908 GAGATATTTTCCTGCAACAAGCACAGGGAATGCCGGAGGGTGGGGAAGGATCACC 967

Qy    261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
        968 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTGATTTGCTACAA 1027

Qy    281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
        1028 CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 1087

Qy    301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
        1088 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTCAGTGCTG 1147

Qy    321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
        1148 TTTATCGCGGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAACTGAACTGG 1207

Qy    341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
        1208 ACGCTTCCCGGTCAGCCGGATAACACGCCGCCGGGAACTGGTGTTTGAACGCTGG 1267
```

```
Qy    361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268 CGTCGGACGTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTGTTCCAGACTTTACAG 1327

Qy    381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACCCCGCCGGGAGAGGTGAAACTGACC 1387

Qy    401 LeuAlaGlyCysGluGlnArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388 CTGGCAGGATGTGAACAGCGTAATGCGCAGGGCATGTGTTCGCTGGCAGGGTTTTACGCAA 1447

Qy    421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
              ||||||||||||||||||||||||||||||||||||
Db   1448 ATCGTGAATGAAGCACGCATACCGGCCTGCAGTTTG 1483
```

RESULT 14
ID   ADC87744 standard; DNA; 1901 BP.
XX
AC   ADC87744;
XX
DT   01-JAN-2004 (first entry)
XX
DE   DNA encoding Escherichia coli appA phytase.
XX
KW   Phytase; food supplement; enzyme delivery matrix; soybean meal;
KW   thermostability; thermotolerance; kernel; phytate;
KW   myo-inositol-hexaphosphate; inositol; inorganic phosphate;
KW   thermotolerant; feed value; digestion; gene; ds; appA.
XX
OS   Escherichia coli.
XX
FH   Key   Location/Qualifiers
FT   CDS   188..1486
FT   /*tag= a
FT   /product= "AppA phytase"
FT   /transl_except= (pos:401..403,aa:Arg)
XX
PN   US2003103958-A1.
XX
PD   05-JUN-2003.
XX
PF   24-MAY-2002; 2002US-00156660.
XX
PR   13-AUG-1997; 97US-0091079S.
PR   01-MAR-1999; 99US-00255214.
PR   13-APR-1999; 99US-00291931.
PR   25-MAY-1999; 99US-00318528.
PR   25-MAY-2000; 2000US-00580515.
PR   24-MAY-2001; 2001US-0086 6379.
XX
PA   (DIVE-) DIVERSA CORP.
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI   Mathur EJ;
XX
DR   WPI; 2003-787039/74.
DR   P-PSDB; ADC87745.
XX
PT   New nucleic acid encoding a polypeptide having phytase activity, useful
PT   in improving the feeding value of phytate rich ingredients or as an aid
PT   in phytate digestion.
XX
PS   Example 1; SEQ ID NO 3; 113pp; English.
XX
CC   The invention discloses a new isolated or recombinant nucleic acid which
CC   encodes a polypeptide having a phytase activity. Also claimed is a
CC   nucleic acid probe, an amplification primer sequence pair, an expression
CC   cassette comprising the nucleic acid, a vector comprising the nucleic
CC   acid, a transgenic non-human animal or plant, or its seed, comprising the
CC   nucleic acid, an antisense oligonucleotide, inhibiting the translation of
CC   a phytase message in a cell, a heterodimer comprising the polypeptide and
CC   a second domain, an array comprising immobilised polypeptide or nucleic
CC   acid, a hybridoma comprising an antibody that specifically binds to the
CC   polypeptide, a food supplement for an animal, an edible enzyme delivery
CC   matrix, an edible pellet comprising a granule edible carrier and the
CC   polypeptide, a feed composition, a soybean meal, isolating or identifying
CC   the polypeptide, making an anti-phytase antibody, producing a recombinant
CC   polypeptide, determining whether a compound binds to the polypeptide,
CC   identifying a modulator, whole cell engineering of new or modified
CC   phenotypes by using real-time metabolic flux analysis, increasing
CC   thermotolerance or thermostability of the phytase polypeptide, increasing
CC   the resistance of the phytase polypeptide to enzymatic inactivation in a
CC   digestive system and processing of corn and sorghum kernels. The phytase
CC   activity comprises catalysis of phytate (myo-inositol-hexaphosphate) to
CC   inositol and inorganic phosphate and is thermotolerant. The nucleic acid
CC   is useful in improving the feeding value of phytate rich ingredients or
CC   as an aid in phytate digestion. The sequence presented is the DNA
CC   encoding the Escherichia coli appA phytase.
XX
SQ   Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:   8.17e-195   Length: 1901
Score: 2182.00   Matches: 424
Percent Similarity: 98.1%   Conservative: 0
Best Local Similarity: 98.1%   Mismatches: 8
Query Match: 96.6%   Indels: 0
DB: 10   Gaps: 0

US-10-601-319-10 (1-432) x ADC87744 (1-1901)

```
Qy      1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188 ATGAAAGCGATCTTAATCCCATTTTTATCCTTGTTAATCCCGTTAACCCCGCAATCTGCA 247

Qy     21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248 TTCGCTCAGAGTGAGCCGGAGCTGAAAGCTGGAAAGTGTGGTGATTGTCAGTGTGTCATGGT 307

Qy     41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308 GTGCGTGCTCCAACCAAAGGCCAACAGCTGATGCAGGATGTCACCCCAGACGCATGGCCA 367

Qy     61 ThrTrpProValLysGluGlyLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    368 ACCTGGCCGGTAAAACTCGGGTTGGCCTGACACCCGCGNGGGTGAGCTAATCGCCTATCTC 427

Qy     81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428 GGACATTACCAACGCCAGCCGCTCTGGTAGCCGACACGGATTGCTCGGCGAAAAAGGCTGCCCG 487

Qy    101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488 CAGTCTGGTCAGGTCGCGGATTATTGCTGACGTGCGACGGCGTACCCGTAAAACAGGCGAA 547

Qy    121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548 GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 607

Qy    141 SerSerProAspProGluPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608 TCCAGTCCCGATCCGGTTAATTAATCCTCTAAAAACTGGCGTTTGCCAACTGCGATAACGCG 667

Qy    161 AsnValThrAspAlaIleLeuGluArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    668 AACGTGACTGACGCGATCCTCAGCGAGGCGGCAGGAGGCTCAATTGCTGACTTTACCGGGCAT 727

Qy    181 TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    728 CGGCAAACGGCGTTTCCGGAACTGGAACGGGTGCTTAATTTCCCGCAATCAAACTTGTGC 787

Qy    201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    788 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC 847
```

```
RESULT 15
ID   AD50297 standard; DNA; 1901 BP.
XX
AC   AD50297;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Escherichia coli K12 appA phytase DNA.
XX
KW   AppA phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW   phytate; animal feed; fish feed; dough; baking; gene; ds.
XX
OS   Escherichia coli; K12.
XX
FH   Key             Location/Qualifiers
FH
FT   CDS             188..1486
FT                   /tag= a
FT                   /product= "AppA phytase"
FT   unsure          401..403
FT                   /tag= b
FT                   /note= "Encodes Arg"
XX
PN   US20040091968-A1.
XX
PD   13-MAY-2004.
XX
PF   20-JUN-2003; 2003US-00601319.
XX
PR   13-AUG-1997;  97US-00910798.
PR   01-MAR-1999;  99US-00259214.
PR   13-APR-1999;  99US-00291931.
PR   25-MAY-1999;  99US-00318528.
PR   25-MAY-2000;  2000US-00580515.
PR   24-MAY-2001;  2001US-00866379.
XX
PA   (SHOR/)  SHORT J M.
PA   (KRET/)  KRETZ K.
PA   (GRAY/)  GRAY K A.
PA   (BART/)  BARTON N R.
PA   (GARR/)  GARRETT J B.
PA   (ODON/)  O'DONOGHUE E.
PA   (MATH/)  MATHER E J.
XX
PI   Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI   Mather EJ;
XX
DR   WPI; 2004-374952/35.
DR   P-PSDB; AD50298.
XX
PT   Producing phytase, involves providing nucleic acid derived from bacteria
PT   encoding polypeptide having phytase activity, and expressing nucleic acid
PT   in yeast.
XX
PS   Disclosure; SEQ ID NO 7; 74pp; English.
XX
CC   The invention relates to a method for producing phytase that involves
CC   providing a nucleic acid encoding phytase derived from a bacteria, and
CC   expressing the nucleic acid in a yeast under conditions that allow
CC   expression of the enzyme in the yeast. The invention also relates to
CC   modified phytase enzyme which has improved thermal tolerance and protease
CC   stability at low pH. The phytase enzyme can be used in foodstuffs to
CC   improve the feeding value of phytate rich ingredients, and in diet of
CC   numerous animals including mammals, fowls and fishes, commercially
CC   significant animals such as pigs, goats, laboratory rodents, commercially
CC   significant avian species such as chicken, ducks, doves, parrot, etc.,
CC   commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC   etc., in dough making and baking, in dietary aids for animals. The method
CC   provides easy manufacture of the active ingredient loaded biocompatible
CC   composition, higher yields and loading efficiency. The phytase
CC   incorporated in to the dietary aid is safe for animals. The present
CC   sequence is Escherichia coli K12 appA phytase DNA.
XX
SQ   Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;
```

Alignment Scores:
Pred. No.:
Score: 2182.00   Matches: 1901
Percent Similarity: 98.1%   Conservative: 424
Best Local Similarity: 98.1%   Mismatches: 8
Query Match: 96.6%   Indels: 0
DB: 12   Gaps: 0

Length: 1901

US-10-601-319-10 (1-432) x AD50297 (1-1901)

```
Qy          1  MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProGlnSerAla    20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188  ATGAAAGCGGATCTTAATCCATTTTTATCTCTTCTGATCCCTCAATCTGCA   247

Qy         21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly    40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248  TTCGCTCAGAGTGAGCCCGAGCTGAAAGCTGGAAAGTGTGGGTGATTGTCAGTCGTCATGGT   307

Qy         41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAlaTrpPro    60
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308  GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGCGTGGCCA   367

Qy         61  ThrTrpProValLysGluGlyLeuThrProArgGlyGluLeuIleAlaTyrLeu    80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368  ACCTGGCCGGTAAAACTGGGTTCGGCTGACACCCCGNGGTGAGCTAATCGGCTATCTC   427
```

Qy   221 LysValSerAlaAspCysValSerLeuThrGlyValAlaValSerLeuAlaSerMetLeuThr   240
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   848 AAGGTGAGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAAATGCTGACG   907

Qy   241 GlyIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr   260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   908 GAGATATTTCTCCTGCAACAAGCACAGGGAAATGCCGGAGCCCGGGGGTGGGAAGGATCACC   967

Qy   261 AspSerHisGlnThrrpAsnThrLeuSerLeuHisAsnAlaGlnPheLeuLeuGln   280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   968 GATTCACACCAGTGGAACACCTTGCTAGGTTTGCATAACGCGCAATTTTATTTGCTACAA   1027

Qy   281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuLysThrAla   300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028 CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTAGATTTGAAGACAGCG   1087

Qy   301 LeuThrProHisProProGlnLyssGlnAlaTyrGlyValThrrleuProThrSerValLeu   320
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGGTGACCATTACCCACTTCAGTGCTG   1147

Qy   321 PheIleAlaGlyHisGluAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp   340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148 TTTATCCGGCGACACGAACTAATTCTGGCGAAACCTCGCGGGCCACTGGAGCTCAACTGG   1207

Qy   341 ThrLeuProGlyGlnProAspAsnThrProGlyGlyGluLeuValPheGluArgTrp   360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208 ACGCTTCCCGGTCAGCCCGGATAAACACCGCCAGGTGGTGAACCTGGTGTTTGAACGCTGG   1267

Qy   361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln   380
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268 CGTCGGCTGAAGCGATAACACAGCCAGTGGATTCAGGTTCGCTGGTCTTCCAGACTTTACAG   1327

Qy   381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProGlyGlyValLysLeuThr   400
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328 CAGATGCGTGATAAAACCCCGCTGTCATTAAATACGCCCGGCGGCGTGAAACTGACC   1387

Qy   401 LeuAlaGlyCysGluGluArgAsnAlaGlyGlyMetCysSerLeuAlaGlyPheThrGln   420
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388 CTGGCAGGATGTGAAGAGCGAAATGCCGGCGGCATGTGTTCGTTGGCAGGTTTTACGCAA   1447

Qy   421 IleValAsnGlnAlaArgIleProAlaCysSerLeu   432
               ||||||||||||||||||||||||||||||||||||
Db   1448 ATCGTGAATGAAGAGCACGCATACCGGCGTGCAGTTTG   1483
```

```
Qy    81  GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro  100
              | ||  |  |  |||| |     ||| | | | |  || ||||  |  ||| |
Db   428  GGACATTACCAACGCCAGCCGTTCTGGGTTAGCCGGACGGGAATTGCTGGCGAAAAGGCTGCCCG  487

Qy   101  GlnSerGlyGlnValAlaIleIleAlaAspValAlaAspGluArgThrArgLysThrGlyGlu  120
              ||| ||| || || | | || | |   ||  |  | |  ||   | ||  |  ||| ||
Db   488  CAGTCTGGTCAGGTCGCCGATTAATTGCTGATGTCGACGAGAGGCTACCCGTAAAACAGGCGAA  547

Qy   121  AlaPheAlaAlaGlyLeuAlaAlaProAspCysBnAlaIleThrValHisGlnAlaAspThr  140
              ||| ||| || ||| |  ||  |  || | |  |  | | |  || | |  ||| |
Db   548  GCCTTCGCCGCCGGCCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG  607

Qy   141  SerSerProAspProLeuAsnProLeuLeuLysThrGlyLysValCysGlnLeuAspAsnAla  160
              ||| ||| || ||  || || ||  | | |  || ||  |  || ||| || ||| ||
Db   608  TCCAGTCCCGATCCGTTATTTAATCCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG  667

Qy   161  AsnValThrAspAlaIleLeuGluArgGlnAlaGlyGlySerIleAlaAlaAspPheThrGlyHis  180
              ||| || || ||  || ||  | | |  ||| || || |  | | ||| || ||| |
Db   668  AACGTGACTGACGCGATCCTCAGCACGCAGGCAGGAGGGTCAATTGCTGCAGATTTACCGGGCAT  727

Qy   181  TyrGlnThrAlaPheArgGluLeuGlyLeuAsnPheProGlnSerAsnLeuCys  200
              | |  || ||| || || ||  |  || ||  || || || ||| ||
Db   728  CGGCAAACGGCGTTTCGGGAACTGGGACTGAACTTTCCGCAATCAAAACTTGTGC  787

Qy   201  LeuLysArgGluLysGlnAspGluSerCysSerSerLeuThrGlnAlaLeuProSerGluLeu  220
              || || ||  | |  ||| |  || | ||  || ||  |  || || || ||
Db   788  CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCAATTACCATCGGAACTC  847

Qy   221  LysValSerAlaAspCysValSerValLeuThrGlyAlaValSerLeuAlaSerMetLeuThr  240
              || || ||| || | |  ||| || || || || ||| || || ||  |  || |
Db   848  AAGGTGAGCGCCGACAAATGTCTCATTAACCGGTGCGGTAAGCCTCCCATCAAATGCTGACG  907

Qy   241  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla  260
              || ||| || || || || || || | | | |  ||  ||  | |  || |
Db   908  CGCACGGCCAGAGGTTGCCCCGAGCCCACCCCGTTATTAGATTTGATCAAGACAGCG  967

Qy   261  GluIlePheLeuLeuGlyIleAlaGlyIleGlyMetProGluProGlyTrpGlyArgIleThr  280
              | | | |  || ||| | ||| | || | |  ||  || | ||  | || || |
Db   968  GAGATATTTCTCCTGGCAATCGCAGGGGATGGGAATGCCCGAGCCCGGATGGGAAGGGATCACC  1027

Qy   281  AspSerHisGlnIleTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln  300
              || ||| || ||  | || |  || || || || ||  | ||  || || || |
Db  1028  GATTCACACACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCCAATTTATTGCTACAA  1087

Qy   301  LeuThrProHisProProGlnLysGlnIleAlaTyrGlyValThrLeuProThrSerValLeu  320
              || || ||  ||  ||| ||  | |  | | | | || || ||  ||  | ||
Db  1088  TTGACGCCCCATCCACCGCAAAAACAGGGGTATGGTGTGACATTACCCACTTCAGTGCTG  1147

Qy   321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340
              | |  || || || ||  | || || || || || || || || || || ||
Db  1148  TTTATCGCCGGACACGATAACTAATCTGGCAAAATCTCCGGCGGCCACTGGAGCTCAACTGG  1207

Qy   341  ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp  360
              || || || ||  || ||  | ||| || || || || || ||  | | ||
Db  1208  ACGCGTTCCCGGTCAGCCGGATAAACACGCCCGCCAGGTGGGTGAACTGGTGTTTGAACGCTGG  1267

Qy   361  ArgArgLeuSerAspAsnSerGlnInTrpIleGlnValSerLeuValPheGlnInThrLeuGln  380
              || || || || ||  | | |  | || |  ||  | || |  |  | |  || |
Db  1268  CGTCGGCCTAAGCGAATAACAGCCAGTTCGGCTTCTTCCAGACTTTACAG  1327

Qy   381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLysLeuThr  400
              || |  || || || || || || || || || ||  | || || || || ||
Db  1328  CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCCGGAGAGGTGAAAACTGACC  1387

Qy   401  LeuAlaGlyCysGluGlyArgAsnAlaGlnInGlyMetCysSerSerLeuAlaGlyPheThrGln  420
              || || || | || || || || ||  | | | |  | ||  || || || ||  |  |
Db  1388  CTGGCAGGGATGTGAACAGCGAAAATGCCCAGGGCATGTGTTCGTTGCGAGGTTTTACGCAA  1447

Qy   421  IleValAsnGluAlaArgIleProAlaCysSerLeu  432
              | |  || | || || | || ||  | || ||
Db  1448  ATCGTGAATGAAGCACGCCATTACCGGGCTGCAGTTTG  1483
```

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

Run on: June 14, 2006, 14:19:34 ; Search time 1412.42 Seconds
                                    (without alignments)
                                    5637.400 Million cell updates/sec

Title:
Perfect score: 2258
Sequence: 1 MKAILIPPLSLLIPLTPQSA............CSLAGPTQIVNEARIPACSL 432

Scoring table:
Sequence: US-10-601-319-10

Scoring table: BLOSUM62
                 Xgapop 10.0 ; Xgapext 0.5
                 Ygapop 10.0 ; Ygapext 0.5
                 Fgapop 6.0 ; Fgapext 7.0
                 Delop   6.0 ; Delext   7.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
  -MODEL=frame+_p2n.model -DEV=xlp
  -Q=/abss/ABSSWEB_spool/US10601319/runat_12062006_180103_27735/app_query.fasta_1
  -DB=Published Applications NA Main -QFMT=fastap -SUFFIX=rnpbm -MINMATCH=0.1
  -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
  -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
  -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
  -MAXLEN=2000000000 -HOST=abss06p
  -USER=US10601319 @CGN 1_3524 @runat_12062006_180103_27735 -NCPU=6 -ICPU=3
  -NO_MAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
  -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
  -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :   Published Applications NA Main:*
    1:    /EMC_Celerra_SIDS3/ptcdata/2/pubpna/US07_PUBCOMB.seq:*
    2:    /EMC_Celerra_SIDS3/ptcdata/2/pubpna/US08_PUBCOMB.seq:*
    3:    /EMC_Celerra_SIDS3/ptcdata/2/pubpna/US09A_PUBCOMB.seq:*
    4:    /EMC_Celerra_SIDS3/ptcdata/2/pubpna/US09B_PUBCOMB.seq:*
    5:    /EMC_Celerra_SIDS3/ptcdata/2/pubpna/US09C_PUBCOMB.seq:*
    6:    /EMC_Celerra_SIDS3/ptcdata/2/pubpna/US10A_PUBCOMB.seq:*
    7:    /EMC_Celerra_SIDS3/ptcdata/2/pubpna/US10B_PUBCOMB.seq:*
    8:    /EMC_Celerra_SIDS3/ptcdata/2/pubpna/US10C_PUBCOMB.seq:*
    9:    /EMC_Celerra_SIDS3/ptcdata/2/pubpna/US10D_PUBCOMB.seq:*
   10:    /EMC_Celerra_SIDS3/ptcdata/2/pubpna/US10E_PUBCOMB.seq:*
   11:    /EMC_Celerra_SIDS3/ptcdata/2/pubpna/US10F_PUBCOMB.seq:*
   12:    /EMC_Celerra_SIDS3/ptcdata/2/pubpna/US10G_PUBCOMB.seq:*
   13:    /EMC_Celerra_SIDS3/ptcdata/2/pubpna/US11A_PUBCOMB.seq:*
   14:    /EMC_Celerra_SIDS3/ptcdata/2/pubpna/US11B_PUBCOMB.seq:*
   15:    /EMC_Celerra_SIDS3/ptcdata/2/pubpna/US11C_PUBCOMB.seq:*
   16:    /EMC_Celerra_SIDS3/ptcdata/2/pubpna/US11D_PUBCOMB.seq:*

SUMMARIES

Result No.	Score	Query % Match	Length	DB ID	Description
1	2258	100.0	1296	14	US-11-056-354-1	Sequence 1, Appli

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

	Score	%			
2	2258	100.0	1308	6	US-10-156-660-1	Sequence 1, Appli
3	2258	100.0	1308	8	US-10-601-319-9	Sequence 9, Appli
4	2258	100.0	1308	10	US-10-601-319-9	Sequence 9, Appli
5	2195	97.2	1901	8	US-10-601-319-6	Sequence 5, Appli
6	2190	97.0	1901	3	US-10-601-319-6	Sequence 6, Appli
7	2190	97.0	1299	8	US-10-601-319-5	Sequence 7, Appli
8	2182	96.6	1901	10	US-10-282-122A-7167	Sequence 7167, Ap
9	2182	96.6	1901	3	US-09-866-379-7	Sequence 7, Appli
10	2182	96.6	1901	3	US-09-866-379-9	Sequence 9, Appli
11	2182	96.6	1901	6	US-10-156-660-3	Sequence 3, Appli
12	2182	96.6	1901	8	US-10-601-319-7	Sequence 7, Appli
13	2182	96.6	1901	14	US-11-056-354-3	Sequence 3, Appli
14	2175	96.3	1901	10	US-10-933-115-7	Sequence 6, Appli
15	2171	96.1	1901	10	US-09-777-566A-1	Sequence 1, Appli
16	2171	96.1	1323	3	US-09-866-379-1	Sequence 1, Appli
17	2171	96.1	1323	3	US-09-866-379-1	Sequence 1, Appli
18	2171	96.1	1323	3	US-10-034-985-1	Sequence 1, Appli
19	2171	96.1	1323	5	US-10-430-356-1	Sequence 1, Appli
20	2171	96.1	1323	6	US-10-601-319-1	Sequence 1, Appli
21	2171	96.1	1323	8	US-10-601-319-1	Sequence 1, Appli
22	2171	96.1	1323	10	US-10-933-115-1	Sequence 1, Appli
23	2168.5	96.0	11357	13	US-11-074-522-14	Sequence 14, Appli
24	2164	95.8	1486	13	US-11-018-709-2	Sequence 2, Appli
25	2157	95.5	1281	7	US-10-282-122A-41608	Sequence 41608, A
26	2157	95.5	1281	7	US-10-334-671-4	Sequence 4, Appli
27	2157	95.5	1489	7	US-10-334-672-4	Sequence 4, Appli
28	2147	95.1	1489	6	US-10-021-723A-1	Sequence 1, Appli
29	2147	95.1	1489	7	US-10-021-723A-9	Sequence 9, Appli
30	2144	95.0	1486	10	US-10-284-962-1	Sequence 1, Appli
31	2144	95.0	1486	13	US-10-282-122A-3181	Sequence 4, Appli
32	1179	52.2	1737	10	US-11-018-709-4	Sequence 4, Appli
33	976.5	43.2	1326	6	US-10-450-763-25689	Sequence 25689, A
34	959.5	42.5	1326	6	US-10-021-723A-11	Sequence 11, Appl
35	954.5	42.3	1325	6	US-10-021-723A-3	Sequence 3, Appli
36	934.5	41.4	1325	6	US-10-021-723A-1	Sequence 41608, A
37	782	34.6	1431	10	US-10-021-723A-9	Sequence 9, Appli
38	723.5	32.0	4050	10	US-10-450-763-386	Sequence 386, App
39	713	31.6	466	10	US-10-282-122A-24247	Sequence 24247, A
40	637	28.2	1264	8	US-10-450-763-24247	Sequence 3181, A
41	637	28.2	1264	10	US-10-450-763-30120	Sequence 30120, A
42	554	24.5	11710	7	US-10-240-689-40	Sequence 40, Appl
43	475.5	21.1	1230	6	US-10-021-723A-7	Sequence 7, Appli
44	462.5	20.5	1266	6	US-10-021-723A-5	Sequence 5, Appli
45	414	18.3	1326	16	US-11-098-686-9338	Sequence 9338, Ap

RESULT 1
US-11-056-354-1

Sequence 1, Application US/11056354
Publication No. US20050246780A1

GENERAL INFORMATION:
  APPLICANT: SHORT, Jay M.
  APPLICANT: KRETZ, Keith
  APPLICANT: GRAY, Kevin A.
  APPLICANT: BARTON, Nelson R.
  APPLICANT: GARRETT, James B.
  APPLICANT: O'DONOGHUE, Eileen
  APPLICANT: MATHUR, Eric J.
  TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM AND METHODS
  TITLE OF INVENTION: FOR MAKING AND USING THEM
  FILE REFERENCE: 564462001803/D1370-9C1
  CURRENT APPLICATION NUMBER: US/11/056,354
  CURRENT FILING DATE: 2005-02-11
  PRIOR APPLICATION NUMBER: US 10/156,660
  PRIOR FILING DATE: 2002-05-24
  PRIOR APPLICATION NUMBER: US 09/866,379
  PRIOR FILING DATE: 2001-05-24
  PRIOR APPLICATION NUMBER: US 09/580,515
  PRIOR FILING DATE: 2000-05-25
  PRIOR APPLICATION NUMBER: US 09/318,528
  PRIOR FILING DATE: 1999-05-25

ALIGNMENTS

```
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.1
;
; SEQ ID NO 1
; LENGTH: 1296
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Modified phytase enzyme
US-11-056-354-1

Alignment Scores:
Pred. No.:
Score:                  3.27e-255
                        2258.00
Percent Similarity:     100.0%         Length:        1296
Best Local Similarity:  100.0%         Matches:       432
Query Match:            100.0%         Conservative:  0
DB:                     14             Mismatches:    0
                                       Indels:        0
                                       Gaps:          0

US-10-601-319-10 (1-432) x US-11-056-354-1 (1-1296)

Qy     1 MetLysAlaIleLeuIlePheProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla  20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAAAGCGATCTTAATCCCATTTTTATCCTCTTCTGATTCCCTTAACCCGCCAATCTGCA  60

Qy    21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly    40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TTCGCTCAGAGTGAGCCCGAGCTGAAACTGGAAAGTGTGGTCATTGTCAGTCGTCATGGT  120

Qy    41 ValArgAlaAlaProThrLysValThrGlnLeuMetGlnAspValThrProAspAlaTrpPro  60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTGCGTGCTGCACCAAAGGCACAACTGATGCAGGATGTCACCCCAGACGCATGGCCA  180

Qy    61 ThrTrpProValLysGlyValGlyLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu    80
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ACCTGGCCGGTAAAACTGGGGTGAGCTGACAACCACGCGGCGGTGGTGAGCTAAATCGCCTATCTC  240

Qy    81 GlyHisTyrTrpProArgGlnArgGlnLeuPheAsnProLeuLysThrGlyValCysGlyPro  100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGACATTACTGGCCGCGTCAGCGTCAGTTAATCCCTCTAAAAACTGGCGTGTGCTGGCCCG  300

Qy   101 GlnSerGlyHisValAlaIleIleGluAlaAspValAlaArgLysThrGlyGlu        120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGTCTGGTCAGGTCGGGATTATTGTCGATGCGGCAGACGCTACCCGTAAAACAGGCGGAA  360

Qy   121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleGluThrValHisThrGlnAlaAspThr  140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAGAGACCGTACATACCCAGGCAGATACG  420

Qy   141 SerSerProAsnProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAspAlaIleAla  160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG  480

Qy   161 AsnValThrAspProAlaIleLeuGluArgGlnAlaGlyGlySerIleAlaAlaAspPheThrGlyHis  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AACGCTGACTGACGCGATCCTCGAGAGGCAGGCAGAGGGTCAATTGCTGACTTTACCGGGCAT  540

Qy   181 TyrGlnThrAlaPheArgGluLeuGlyValLeuAsnProProGlnSerAsnLeuCys      200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TATCAAACGGCCGTTTCCGGACTGGCTGGTGCTTAATTTTCCGGGAATCAAACTTGTGC  600

Qy   201 LeuLysGluArgGluLysArgGlnInAspSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTTAAACGTGAGAAAACAGGACGCGAAGCTGTTCATTAACGCAGGCAATTACCATCGGAACTC  660

Qy   221 LysValLeuSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AAGGTGAGCGCGGCCGACTGTGTCTCATTAACCGGTGCGGTGAGCCTCGCCATCAATGCTGACG  720
```

```
Qy   241 GluIlePheLeuLeuGlnIleAlaGlnGlyMetProGluProGlyTrpGlyTyrGlyIleThr  260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GAGATATTTTCTTCTGCAACAGCACAGGGAATGCCGGAACCGGGTTGGGGATACGGGATCACC  780

Qy   261 AspSerHisGlnThrAlaPheThrLeuLeuSerLeuHisAsnAlaGlnInPheAspLeuLeuGln  280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GATTCACACCAGTGGACACCTTGCTAAGTTTGCATAACGGCGCAATTGATTTGCTACAA  840

Qy   281 ArgThrProGlyValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CGCACGCCAGGAGGTTGCCCGCAGCCGCCACCCCGTTATTAGATTTGATCAAGACAGCG  900

Qy   301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu   320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTCAGTGCTG  960

Qy   321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp   340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TTTATCGCCGGACACGATACTAATCTCGCGAATCTCGGCGGTGCTTTAGAGCTCAACTGG  1020

Qy   341 ThrLeuProGlyGlnProTyrAspAsnThrProProGlyGlyGlyLeuValPheGluLeuTrp  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 ACGCTTCCCGGTCAGCCCCGGTAACACCGCGCCAGGTGGTGGAACTGGTGTTTGAACGCTGG  1080

Qy   361 ArgArgLeuSerAspAsnSerGlnInTrpIleGluValSerLeuValPheGlnThrLeuGln  380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTGAAGTTCGGCTGGTCTTCCAGACTTTACAG  1140

Qy   381 GlnMetArgAspProLeuSerLeuAsnThrProProGlyGluValLysLeuThr        400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 CAGATGCGTGATAAAACGCCCGCTGTCATTAAATACGGCGCCCGGAAGAGGTGAAACTGACC  1200

Qy   401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln   420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 CTGGCAGGATGTGAAGAGCGAAATGCCCAGGGCATGTGTTCGTTGGCAGGTTTTACGCAA  1260

Qy   421 IleValAsnGluAlaArgIleProAlaCysSerLeu  432
          ||||||||||||||||||||||||||||||||||||
Db  1261 ATCGTGAATGAAGCACCCATACCGGCGTGCAGTTTG  1296
```

```
;
; LENGTH: 1308
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified phytase enzyme
; FEATURE:
; NAME/KEY: CDS
US-10-156-660-1  LOCATION: (1)...(1308)
; Query Match:
DB:

Alignment Scores:
Pred. No.:
Score:                  3.32e-255
Score:                  2258.00
Percent Similarity:     100.0%
Percent Identity:       100.0%     Matches:        1308
Best Local Similarity:  100.0%     Matches:         432
Query Match:            6          Conservative:      0
DB:                                Mismatches:        0
                                   Indels:            0
                                   Gaps:              0
                                   Length:         1308
                                   Length:          432

US-10-601-319-10 (1-432) x US-10-156-660-1 (1-1308)

Qy   1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuLeuIleProLeuThrProGlnSerAla 20
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAGCGGATCTTAATCCCATTTTATCTCTTCTTGAATCCGCTTAACCCCGCAAATCTGCA 60

Qy  21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleSerArgHisGly 40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTAGTCAGTCGTCATGGT 120

Qy  41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 GTGCGTGCGCCTACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy  61 ThrTrpProValLeuGlyLeuGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 ACCTGGCCGGTAAAACTGGGTGAGCTGACACCCCGCGGTGGTGAGCTAATCGCCTATCTC 240

Qy  81 GlyHisTyrTrpArgGlnAsnArgLeuValAlaAlaGlyLeuLeuProLysCysGlyCysPro 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 GGACATTACTGGCGTCAGGTCGGTCGCGGAATTGCTCGCTGCCTAAAATGTGGCTGCCCG 300

Qy 101 GlnSerGlyIleValAlaIleLeuIleGlnAlaAspValAlaSerGlyThrArgLysThrGlu 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 CAGTCTCGGTCAGGTCGCGGATTATGTCGAGCACGAGGCTACCGTAAAACAGGGCGAA 360

Qy 121 AlaPheAlaAlaGlyLeuAlaProGlyLeuAlaIleGluThrValHisThrGlnAlaAspThr 140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 GCCTTCGCGCCGGGCTGGCCACTGACTGTGCAATAACCGTACACTACCCAGGCAGATACG 420

Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyThrGlyLeuAspAlaAsnAla 160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 TCCAGTCCGATCCGTTATTTAATCCTCTAAAAACTGGCTTTGCCAACTGGATACCGCCG 480

Qy 161 AsnValThrAspAlaIleLeuGluLeuGluArgAlaGlyAlaGlySerIleLeuAlaAspPhePheThrGlyHis 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 AACGTGACTGACGCGATCCTCGAGGGCGCAGGGGTCAATTGCTGACTTTACCGGGCCAT 540

Qy 181 TyrGlnThrLeuGlnPheArgGluLeuGluLeuValGlnValLeuAsnAsnProGlnSerAsnLeuCys 200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 TATCAAACGGGGTTTCCGGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600

Qy 201 LeuLysArgGluLysGlyAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCTCGGAACTGGAACGGGTGCTTAATTTTCCG 660

Qy 221 LysValSerAlaAspCysValSerLeuThrArgAlaValAlaValLeuSerMetLeuThr 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 661 AAGGTGAGCGCCGACTGTGTCTCATTAACGCGGGCTGTGGCCGTCCTGAGCATGCTGACG 720

Qy 241 GluIlePheLeuLeuGlnIleAsnGlyMetProGluProGlyTrpGlyArgIleThr 260
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 721 GAGATATTTCTCCTGCAACAGATTAACGGGATGCCGGAGCCGGGGTGGGGGAGGATCACC 780

Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
```

RESULT 3
US-10-601-319-9
; Sequence 9, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING THEM
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 9
; LENGTH: 1308
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified phytase enzyme
; FEATURE:

```
Db 781 GATTCACACAGTGGAACACCTTGCTTAAGTTGCATAACCGCAATTTGATTGTACACA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 281 ArgThrProGluValAlaAlaSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 841 CGCACGCCAGAGGTTGCCGCAAGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 900

Qy 301 LeuThrProHisProProGlyLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 901 TTGACGCCCCATCCACCGCAAAAACAGGCCTATGGTGTGACATTACCAACTTCAGTGCTG 960

Qy 321 ArgArgLeuSerAspAsnSerGlnThrTrpIleGlnValSerLeuValPheGlnThrLeuGln 340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 961 CGTCGGCTAAGCGATAACAGCCAGTGGAATCCAGGTTTCGGCTGGGCGTCTTCCAGACTTTACAG 1020

Qy 361 GlnMetArgAspAspLysThrProLeuSerLeuAsnThrProProGlyGlyGluValLeuLysLeuThr 400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 CAGATGCGTGATAAAACGCCGCTCTCATTAAATACGCCGCCCGGAGGTGAAACTGACC 1080

Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnIleGlyMetCysSerSerLeuAlaGlyPheThrGln 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGCAATTGGTATGTGTTCGTCTTTACGGCAA 1260

Qy 421 IleValAsnGluValAlaArgIleProAlaCysSerLeu 432
         ||||||||||||||||||||||||||||||||||||||
Db 1261 ATCGTTGAATGAAGCACGCATACCGGCGTGCAGTTTG 1296
```

```
;          NAME/KEY: CDS
;          LOCATION: (1)...(1296)
US-10-601-319-9

Alignment Scores:
Pred. No.:
Score:                    3258.00     Length:        1308
Percent Similarity:       100.0%      Matches:       432
Percent Identity:         100.0%      Conservative:  0
Best Local Similarity:    100.0%      Mismatches:    0
Query Match:              8           Indels:        0
DB:                                   Gaps:          0
```

```
US-10-601-319-10 (1-432) x US-10-601-319-9 (1-1308)

Qy    1  MetLysAlaIleIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla   20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA     60

Qy   21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleSerArgHisGly        40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  TTCGCTCAGAGTGAGCCGGAGCTGAAAGTGGAAAGTGTCGTGATTTCAGTCGTCANGGT     120

Qy   41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro     60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA    180

Qy   61  ThrTrpProValLeuGlyLeuGlyIleThrProArgGlyIleGluLeuIleAlaTyrLeu     80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  ACCTGGCCGGTAAAACTGGGTGAGCTAACACCGCGGGCGGTGGAGCTTATTGCCTATCTC    240

Qy   81  GlyTyrHisTyrThrProArgGlnIleArgLeuValAlaAlaGlyLeuLeuProLysGly    100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  GGACATTACTACGCCGTCAGCCGTCAGGGTTGAGCCGACGGATTCGCCTAAAGTGCTGCCG    300

Qy  101  GlnSerGlyLeuValAlaIleLeuIleAlaAspValAlaAspGluArgThrArgLysThrGlu  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  CAGTCTGGTCAGGTCCGATATTTGTGATCGACGACGCACTACCCGTAAAACAGCGGAA     360

Qy  121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr    140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  GCCTTCGCCGCCGGTCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG    420

Qy  141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla    160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  TCCAGTCCCGATCCGTTATTTAATCCTCTAAAACTGGCGTTGCCAACTGGATAACGCG     480

Qy  161  AsnValThrAspAlaIleLeuGluArgAlaGlyGlySerIleAlaAspPheThrGlyHis    180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  AACGTGACTGACGCGATCCTCGAGAGGGCAGGAGGTCAATTGCTGACTTTACCGGGCAT    540

Qy  181  TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys    200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  TATCAAACGGCCGGTTTCGGGAACTGGAACGGGCTGAATTTTCCGCAATCAAACTTGTGC    600

Qy  201  LeuLysArgGluGluLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  CTTAAACGTGAGAAAACAGGACCAAAGCTGTTCATTAACCAGCATTACCATCGGAACTC    660

Qy  221  LysValSerAlaAspProCysValSerLeuThrArgGlyAlaValSerLeuAlaSerMetLeuThr  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  AAGGTGAGCGCCGACTGTGTCATTAACCCGTGCCGTCCGCATCAAGTGCTGACG     720

Qy  241  GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr    260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  GAGATATTTCTCCTGCAACAGGCACAGGGAATGCCGGAGCCGGGGTGGGGAGGGATCACC    780

Qy  261  AspSerHisGlnThrTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln  280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  GATTCACACCAGTGGAACACACTTGCTAAGTTTGCATAACGCCCAATTGATTTGCTACAA    840

Qy  281  ArgThrProGlyValAlaArgSerArgAlaThrProLeuLeuAspLeuLeuIleLysThrAla  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  CGCACCCCAGAGGGTTGCCCGCAGCCGCCGGCCACCCCGTTATTAGATTTGATCAAGACAGCG    900
```

```
Qy  301  LeuThrProHisProProGlnHisLysGlnAlaTyrGlyValThrLeuProThrSerValLeu    320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  TTGACGCCGCCACCACCGCAACATAAACAGGCGTATGGGGTGACAATTACCCCACTGTCTTG    960

Qy  321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp    340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG   1020

Qy  341  ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp    360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021  ACGCTTCCCGGTCAGCCCGGATAACACGCCGCCGGGCGGTGAACTGGTGTTTGAACGCTGG   1080

Qy  361  ArgArgLeuSerAspAsnSerGlnThrTrpIleGluGlnValSerLeuValPheGlnThrLeuGln  380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081  CGTCGGCTAAGCGATAACAGCCAGACGTGGATTGAGCAGGTTTCGCTGGTCTTCCAGACTTTACAG  1140

Qy  381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr    400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141  CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCCGGAGAGGTGAAACTGACC   1200

Qy  401  LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln    420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201  CTGGCAGGATGTGAAGAGCGAAATGCCCAGGGCATGTGTTCGCTGGCAGGGTTTACGCAA   1260

Qy  421  IleValAsnGluAlaArgIleLeuProAlaCysSerLeu                          432
         ||||||||||||||||||||||||||||||||
Db 1261  ATCGTGAATGAAGCACCCATACCGGCGTGCAGTTTG                            1296
```

```
Alignment Scores:
Pred. No.:                3.32e-255          Length:   1308
```

```
Score:                    2258.00        Matches:          432
Percent Similarity:       100.0%         Conservative:       0
Best Local Similarity:    100.0%         Mismatches:         0
Query Match:              100.0%         Indels:             0
DB:                       10             Gaps:               0

US-10-601-319-10 (1-432) x US-10-933-115-9 (1-1308)

Qy      1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATCCCGTTAACCCCGCAATCTGCA 60

Qy     21 PheAlaGlnSerGluProGlnLeuLysLeuGluSerValValIleValSerArgHisGly 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCGCTCAGAGTGAGCCGCAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120

Qy     41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGCGTGCGCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy     61 ThrTrpProValGlyGlyLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACCTGGCCGGTAAAACTGGGTGAGCTGACACCGCGTGGTGGTGAGCTAATCGCCTATCTC 240

Qy     81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysGlyCysPro 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGACATTACTGGCGTCAGCGTCTGGTAGCCGACGGATTGCTGCCTAAAGGTTGTCCG 300

Qy    101 GlnSerGlyGlnValAlaIleIleIleAlaAspGlyLeuThrArgGlyLysThrGlyGlu 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGTCTGGTCAGGTCGCGATTATTATTGCTGATGGCCTGACGCGTGGTAAAACCGGCGAA 360

Qy    121 AlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCTTCGCCGCCGGCCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420

Qy    141 SerSerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG 480

Qy    161 AsnValThrAspAlaIleLeuGluArgAlaGlyAspGlySerIleAlaAspPheThrGlyHis 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AACGTGACTGACGCGATCCTCGAGAGGGCCAGGAGGGTCAATTGCTACTTTACCGGGCAT 540

Qy    181 TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TATCAAACGGCGTTTCGGGAACTGGAACGGGTGCTTAATTTCCGCAATCAAAACTTGTGC 600

Qy    201 LeuLysArgGlyLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATGGGAACTC 660

Qy    221 LysValSerArgAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AAGGTGAGCGCCGACTGTGTCTCATTAACCGGTGCGGTAAGCCTCGCAATCAATGCTGACG 720

Qy    241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCCGAGCCGGGGGAGGGATCACC 780

Qy    261 AspSerHisGlnThrLeuAsnAlaGlnPheAspLeuLeuGln 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GATTCACACCAGTGGACACCTTGCTAAGTTTCCATACGCCCAATTTGATTTGCTACAA 840

Qy    281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 900

Qy    301 LeuThrProHisProProGlnLeuSerGlnValThrGlyValThrLeuProThrSerValLeu 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACAATACCCACCTTCAGTGTCT 960

Qy    321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    321 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACAATACCCACCTTCAGTGTCT 340
```

```
Db    961 TTTACGCCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGGCCGGACTGGAGCTCAACTGG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    341 ThrLeuProGlyGlnProAlaAspAsnThrProGlyGlyGlyLeuGluLeuValPheGluArgTrp 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGGTTCCCCGGTCAGCCCGGATAACACGCCCCCGGTGGGTGGACTGG 1080

Qy    361 ArgArgLeuSerArgAspAsnSerGlnTrpIleGlnIleValSerLeuSerLeuValPheGlnThrLeuGln 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CGTCGGCTAAGCGATAAACAGCCAGTTCAGATTGTTTCGCTGTCTCTGGTCTTCCAGACTTTACAG 1140

Qy    381 GlnMetArgAlaAspProLeuProLeuSerLeuAsnThrProGlyGlyGlyLeuLysLeuValPheGluArgThr 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAGATGCGTGCATGAAAAACGCGCGTCATTAAAATACGCCCGCCGGAGAGGTGAAAACTGACC 1200

Qy    401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerSerLeuValAlaGlyPheThrGln 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGCAGGATGTGAAGAGCGCAAAATGCGCAGGGCAATGTGTTCGTTGGGCAGGTTTTTACGCAA 1260

Qy    421 IleValAlaAsnGluAlaArgIleProAlaCysSerLeu 432
          |||||||||||||||||||||||||||||||||||||||
Db   1261 ATCGTGAATGAAGCACGCGATACCCGGCGTGCAGTTTG 1296
```

```
RESULT 5
US-10-601-319-6

; Sequence 6, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
;  APPLICANT: Short, Jay M.
;  APPLICANT: Kretz, Keith A.
;  APPLICANT: Gray, Kevin A.
;  APPLICANT: Barton, Nelson Robert
;  APPLICANT: Garrett, James B.
;  APPLICANT: O' Donoghue, Eileen
;  APPLICANT: Mathur, Eric J.
;  TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
;  TITLE OF INVENTION:  AND USING THEM
;  FILE REFERENCE: 09010-029011
;  CURRENT APPLICATION NUMBER: US/10/601,319
;  CURRENT FILING DATE:  2003-06-20
;  PRIOR APPLICATION NUMBER: US 09/866,379
;  PRIOR FILING DATE: 2001-05-24
;  PRIOR APPLICATION NUMBER: US 09/580,515
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: US 09/318,528
;  PRIOR FILING DATE: 1999-05-25
;  PRIOR APPLICATION NUMBER: US 09/291,931
;  PRIOR FILING DATE: 1999-04-13
;  PRIOR APPLICATION NUMBER: US 09/259,214
;  PRIOR FILING DATE: 1999-03-01
;  PRIOR APPLICATION NUMBER: US 08/910,798
;  PRIOR FILING DATE: 1997-08-13
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: FastSEQ for Windows Version 4.0
US-10-601-319-6

;  SEQ ID NO 6
;  LENGTH: 1901
;  TYPE: DNA
;  ORGANISM: Escherichia coli
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: 403
;  OTHER INFORMATION: n = A,T,C or G
US-10-601-319-6

Alignment Scores:
Pred. No.:                              1.53e-247     Length:          1901
Score:                                  2195.00       Matches:          425
Percent Similarity:                     98.4%         Conservative:       0
Best Local Similarity:                  98.4%         Mismatches:         7
Query Match:                            97.2%         Indels:             0
DB:                                     8             Gaps:               0

US-10-601-319-10 (1-432) x US-10-601-319-6 (1-1901)
```

RESULT 6
US-10-601-319-5

```
; Sequence 5, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
;  APPLICANT: Short, Jay M.
;  APPLICANT: Kretz, Keith A.
;  APPLICANT: Gray, Kevin A.
;  APPLICANT: Barton, Nelson Robert
;  APPLICANT: Garrett, James B.
;  APPLICANT: O' Donoghue, Eileen
;  APPLICANT: Mathur, Eric J.
;  TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
;  TITLE OF INVENTION: AND USING THEM
;  FILE REFERENCE: 09010-029011
;  CURRENT APPLICATION NUMBER: US/10/601,319
;  CURRENT FILING DATE: 2003-06-20
;  PRIOR APPLICATION NUMBER: US 09/866,379
;  PRIOR FILING DATE: 2001-05-24
;  PRIOR APPLICATION NUMBER: US 09/580,515
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: US 09/318,528
;  PRIOR FILING DATE: 1999-05-25
;  PRIOR APPLICATION NUMBER: US 09/291,931
;  PRIOR FILING DATE: 1999-04-13
;  PRIOR APPLICATION NUMBER: US 09/259,214
;  PRIOR FILING DATE: 1999-03-01
;  PRIOR APPLICATION NUMBER: US 08/910,798
;  PRIOR FILING DATE: 1997-08-13
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;  LENGTH: 1901
;  TYPE: DNA
;  ORGANISM: Escherichia coli
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: 403
;  OTHER INFORMATION: n = A,T,C or G
; US-10-601-319-5
```

Alignment Scores:
Pred. No.:
Score: 5.94e-247   Length: 1901
Percent Similarity: 2190.00   Matches: 425
Best Local Similarity: 98.4%   Conservative: 0
Query Match: 98.4%   Mismatches: 7
DB: 97.0%   Indels: 0
                          8   Gaps: 0

US-10-601-319-10 (1-432) x US-10-601-319-5 (1-1901)

```
Qy    1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
        ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAAATCTGCA

Qy   21 PheAlaGlnSerGluProGlyLeuGlyLeuGluSerValValIleValSerArgHisGly 40
Db  188 PheAlaGlnSerGluProGlyLeuGlyLeuGluSerValValIleValSerArgHisGly 247
        TTCGCTCAGAGTGAGCCGGGACTGGAAAGTGTGGTAATTGTCAGTCGTCATGGT 307

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db  248 ... 367

Qy   61 ThrTrpProValLeuGlyLeuGlyLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db  308 ... 427

Qy   81 GlyHisTyrTrpProGlnGlnArgGlnLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
Db  368 ... 487

Qy  101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db  428 ... 547

Qy  121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db  488 ... 607

Qy  141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db  548 ... 667

Qy  161 AsnValThrArgAlaIleLeuGluAlaAlaGlyGlySerIleAlaAlaPheThrGlyHis 180
Db  608 ... 727

Qy  181 TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db  668 ... 787

Qy  201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db  728 ... 847

Qy  221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db  788 ... 907

Qy  241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db  848 ... 967

Qy  261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
Db  908 ... 1027

Qy  281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
Db  968 ... 1087

Qy  301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 1028 ... 1147

Qy  321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 1088 ... 1207

Qy  341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1148 ... 1267

Qy  361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1208 ... 1327

Qy  381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1268 ... 1387

Qy  401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1328 ... 1447

Qy  421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1388 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTGAATGAAGCACCGCATACCCGGCGTGCAGTTTG 1483
```

```
Qy    248 TTCGCTCAGAGTGAGCCCGGAGCTGGAAAGTGTGGGATTGATTGTGCTTCATGGT 307
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    401 ...                                                      420

Qy    41  ValAlaArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db    308 GTGCGGTTGCTCCAACCAAGGCCACGCAGCTAATGATGCAGGATGTCACCCCAGACGCATGGCCA 367

Qy    61  ThrTrpProValLysLeuGlyLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db    368 ACCTGGCCGGTAAAACTGGGTCTGACACCGGGTGGCTAAATCGGCCTATCTC 427
```

Alignment Scores:
Pred. No.:
Score: 5.94e-247
                                      2190.00
Percent Similarity: 98.4%   Matches: 425
Best Local Similarity: 98.4%   Conservative: 0
Query Match: 97.0%   Mismatches: 7
DB: 10   Indels: 0
                                      Length: 1901
                                      Gaps: 0

US-10-601-319-10 (1-432) x US-10-933-115-5 (1-1901)

```
Qy    1   MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db    188 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 247

Qy    21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db    248 TTCGCTCAGAGTGAGCCCGGAGCTGGAAAGTGTGGGATTGATTGTGCTTCATGGT 307

Qy    41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db    308 GTGCGGTTGCTCCAACCAAGGCCACGCAGCTAATGATGCAGGATGTCACCCCAGACGCATGGCCA 367

Qy    61  ThrTrpProValLysLeuGlyLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db    61  ...                                                      80
```

RESULT 8

US-10-282-122A-7167

Sequence 7167, Application US/10282122A
Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: Patentin version 3.1

SEQ ID NO 7167
  LENGTH: 1299
  TYPE: DNA
  ORGANISM: Escherichia coli

US-10-282-122A-7167

US-10-601-319-10 (1-432) x US-10-282-122A-7167 (1-1299)

Alignment Scores:
Pred. No.:
Score: 2182.00
Percent Similarity: 98.1%
Best Local Similarity: 98.1%
Query Match: 96.6%

Length:	1299
Matches:	424
Mismatches:	8
Conservative:	0
Indels:	0
Gaps:	0

DB:

Qy   1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db   1 ATGAAAGCGGATCTTAATCCATTTTTATCTCTTCTGATTCCGCTAACCCCGCAATCTGCA 60

Qy   21 PheAlaGlnSerGluProGluLeuLysGluSerValValIleValSerArgHisGly 40
Db   61 TTCGCTCAGAGTGAGCCCGAGCTTAAAGAGTCTGGAAAGTGTGGGTGGATTGTCCGTGGT 120

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db   121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

RESULT 9
US-09-866-379-7
; Sequence 7, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 7
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-7

US-10-601-319-10 (1-432) x US-09-866-379-7 (1-1901)

Alignment Scores:
Pred. No.:
Score: 2182.00
                                                    Length: 1901
Percent Similarity: 98.1%   5.22e-246   Matches: 424
Best Local Similarity: 98.1%   Conservative: 0
Query Match: 96.6%   Mismatches: 8
DB: 3   Indels: 8
                                                     Gaps: 0

```
Qy    1 MetLysAlaIleLeuIleIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla  20
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  188 ATGAAAGCGGATCTTAATCCCATTTTATCTCTTCTTGATTCCGGTTAACCCGCAATCTGCA  247

Qy   21 PheLeuAlaGlnSerGlyGluProGluLeuLysLeuGluSerValValIleSerArgHisGly  40
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  248 TTCGCTCAGAGTGAGCCCGAGCTGAAAGTTGGAAAGTGTGGGATAAATTGTCAGTCGTCATGGT  307

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA  367

Qy   61 ThrTrpProValLysLeuGlyLeuGlyTyrGluThrProArgGlyTyrGlyGluLeuIleAlaTyrLeu  80
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  368 ACCTGGCCGGTAAAACTGGGTTTGGGCTA...GAGACACCCGCGGGGTAGTCGGCTAATCTGCA  427

Qy   81 GlyHisTyrTrpArgGlnArgLeuValAlaAlaAspGlyLeuLeuProLysCysGlyCysPro  100
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  428 GGACATTACCAACGCCAGCGTCTGGCTAGCCGACGGATTGCTGCCGAAAAAGGGCTGCCCG  487

Qy  101 GlnSerGlyGlnValAlaIleIleIleAlaAspValAlaAspGlyLeuLeuArgThrArgGlyGlu  120
```

```
Qy   61 ThrTrpProValLysLeuGlyLeuGlyTyrGluThrProArgGlyTyrGlyGluLeuIleAlaTyrLeu  80
Db  181 ACCTGGCCGGTAAAACTGGGTTTGGGCTA...GAGACACCCGCGGGGTATGGTGAGCTGATTGCCTATCTC  240

Qy   81 GlyHisTyrTrpArgGlnArgLeuValAlaAlaAspGlyLeuLeuProLysCysGlyCysPro  100
Db  241 GGACATTACCAACGCCAGCGTCTGGCTAGCCGACGGATTGCTGCCGAAAAAGGGCTGCCCG  300

Qy  101 GlnSerGlyGlnValAlaIleIleIleAlaAspValAlaAspGlyLeuLeuArgThrArgGlyGlu  120
Db  301 CAGTCTGGTCAGGTCGCGATTATTATTGCTGATGTCGACGGGGTTAAAACCAGGGGA  360

Qy  121 AlaPheHisAlaGlyLeuAlaProAspAspCysAlaIleIleThrArgValHisThrGlnAlaAspThr  140
Db  361 GCCTTCCACGCCGGGCTGGCACCTGACGTGTGCAATACCGTACATACCCAGGCAGATACG  420

Qy  141 SerSerProAspProLeuProPheAsnProLeuValSerGlyGlnLeuAspAsnAla  160
Db  421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACGGGGTTTGCCAACTGGATAACGCG  480

Qy  161 AsnValThrAspAlaIleLeuLeuArgAlaGlyGlySerIleIleAlaAspProPheThrGlyHis  180
Db  481 AACGTGACTGACCGGATCCTCACCAGGGCGGGTGGCTCAATTGCTGACTTTACCGGGCAT  540

Qy  181 TyrGlnThrAlaPheAlaArgGluLeuGluValValLeuAsnPheProGlnSerAsnLeuCys  200
Db  541 CGGCAAACGGCGTTTGCGAACGGAGCAGGGTCGTTAATTTTCCGCAATCAAACTTGTGC  600

Qy  201 LeuLysSerArgGlnLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220
Db  601 CTTAAAAGCTGAGAAAACAGGACGAAAGCTGTTCATTAACGCAGGCCATTACCATCGGAACTC  660

Qy  221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr  240
Db  661 AAGGTGAACGCCGACAATGTCTCATTAACCGGTGCGGTTAAGCCTTCGCATCAANGCTGACG  720

Qy  241 GluIlePheLeuLeuGlnAlaGlnIleGlyMetProGlyProGlyTyrThrGlyArgIleThr  260
Db  721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCCGGCCGTGGGGAAGGATACC  780

Qy  261 AspSerHisHisGlnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln  280
Db  781 GATTCACACCACCAGGGAACACACTTGCTAAGTTTGCATAACGCGCAATTTATTTGCTACAA  840

Qy  281 ArgThrProGluValValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla  300
Db  841 CGCACCGCCAGAGGTTGCCCGCAGCCGCGCCACCCGTTATTAGATTTGATCAAGACAGCG  900

Qy  301 LeuThrProHisProProGlnLysGlnIleAlaThrTyrGlyTyrLeuProThrSerValLeu  320
Db  901 TTGACCCCCCATCCACCGCAAAAACAGGGCTATGGGTGTTGACATTACCCACTTCAGTGCTG  960

Qy  321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340
Db  961 TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCCACTGGAGCTGAACTGG  1020

Qy  341 ThrLeuProGlyTyrGlnProProGlyTyrGluLeuValPheGluGlyTrp  360
Db 1021 ACGCTTCCCGGTCAGCCGCCGGATAACACGCCCGGCCAGGTGGGTGAACTGGTGGG  1080

Qy  361 ArgArgLeuSerArgAsnSerGlnGlnThrProIleGluGlnValSerLeuValAlaLeuPheGln  380
Db 1081 CGTCGGCTAACGGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCGCTGGACTTTACAG  1140

Qy  381 GlnMetArgAspLysSerThrProLeuSerLeuAsnThrProProGlyValAlaLysSerLeuThr  400
Db 1141 CAGATGCGTGATAAAAGCACGCCGCTGTCATTAAATACGCCGCCCGGTGTTGCTAAAACTGACC  1200

Qy  401 LeuAlaGlyLysCysGluLeuArgAsnLeuGlyMetCysSerLeuValAlaGlyPheThrGln  420
Db 1201 CTGGCAGGATGTGAAGAGGGAAATGCCAGGGCATGTGTTCGTTGGCAGGTTTTACCGCAA  1260

Qy  421 IleValAlaAsnGluValAlaArgIleProAlaAlaCysSerSerIle  432
Db 1261 ATCGTGAATGAAGCACGCAACACCCGGCGTTCCAGTTTG  1296
```

; APPLICANT: DIVERSA CORPORATION
; GENERAL INFORMATION:
; Patent No. US20020136754A1
; Sequence No. 9, Application US/09866379
US-09-866-379-9
RESULT 10

; APPLICANT: DIVERSA CORPORATION
; GENERAL INFORMATION:
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen

; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1

; SEQ ID NO 9
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-9

Alignment Scores:
Pred. No.:
Score: 5.22e-246
                          2182.00
Percent Similarity: 98.1%   Matches: 424
Best Local Similarity: 98.1%   Conservative: 0   Mismatches: 8
Query Match: 96.6%   Indels: 0   Gaps: 0
DB: 3

US-10-601-319-10 (1-432) x US-09-866-379-9 (1-1901)

Qy   1 MetLysAlaIleLeuIleLeuProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 188 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 247

Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 248 TTCGCTCAGAGTGAGCCGGAGCTGAAACTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 307

Qy   41 ValArgAlaProThrLysMetThrGlnAlaMetArgAspValThrProAspAlaTrpPro 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 367

Qy   61 ThrTrpProValLysGluLeuGlyLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 368 ACCTGGCCGGTAAAAGAGCTGGGTCTGACACCGCGNGGTGGTGAGCTAATCGCCTACTC 427

Qy   81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysGlyGlyCysPro 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 428 GGACATTACCAACGCCAGGCGGTCTGGCCGAAAGGGCTCTGCCCG 487

Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 488 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGGGAA 547

Qy 121 AlaPheAlaAlaGlyTyrLeuAlaProAspCysAlaIleIleThrValHisThrGlnAlaAspThr 140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 548 GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 607

Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 608 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGGGTTTGCCAACTGGATAACGCG 667

Qy   ...sequence alignment continues...

APPLICANT: Short, Jay M.
APPLICANT: Kretz, Keith
APPLICANT: Gray, Kevin A.
APPLICANT: Barton, Nelson R.
APPLICANT: Garrett, James B.
APPLICANT: O'Donoghue, Eileen
APPLICANT: Mathur, Eric J.

TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM
TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
FILE REFERENCE: 09010-029007
CURRENT APPLICATION NUMBER: US/10/156,660

CURRENT FILING DATE: 2002-10-01
PRIOR APPLICATION NUMBER: US 09/866,379
PRIOR FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 09/580,515
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 3
LENGTH: 1901
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: CDS
LOCATION: (188)...(1483)
FEATURE:
NAME/KEY: misc_feature
LOCATION: 403
OTHER INFORMATION: n = A,T,C or G
US-10-156-660-3

US-10-601-319-10 (1-432) x US-10-156-660-3 (1-1901)

Alignment Scores:
Pred. No.:
Score: 5.22e-246   Length: 1901
                             2182.00   Matches: 424
Percent Similarity: 98.1%   Conservative: 0
Best Local Similarity: 98.1%   Mismatches: 8
Query Match: 96.6%   Indels: 0
DB: 6   Gaps: 0

Qy   1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla   20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 ATGAAAGCGGATCTTAAATCCCATTTTATCTCTTCTGATTCCGTTAACCCGCAAATCTGCA 247

Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly   40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248 TTCCCTCAGAGTGAGCCCGAGCTGAAAGCTGGAAAGTGTGGTTGGATTGTCAGTTCGTCATGGT 307

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308 GTGCGTGCTTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCGTGGCCA 367

Qy   61 ThrTrpProValLysGluLeuGlyLeuThrProArgGlyGlyGluLeuLeuIleAlaTyrLeu   80
           |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   368 ACCTGGCCGGTAAAACTCGGTCGGCTGACACCCCGNGGTGGGTGAGCTAATCGCCTATCTC 427

Qy   81 GlyHisTyrTrpArgGlnAsnGlyLeuValAlaAspGlyLeuLeuProLysGlyCysPro 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 GGACATTACCAACGGCCAGGTCTGCTGGTGAGGGACGGGATTGCTGGGCGAAAAAGGCGTGCCG 487

Qy   101 GlnSerGlyGlnValAlaIleAlaIleAlaAlaAspValAspGluArgThrArgLysThrGlyGlu 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488 CAGTCTGGTCAGGTCGGCGGATTAATTGCTGATGTCGACGAGCGTACCCGTAAAAACAGGCGAA 547

Qy   121 AlaPheAlaAlaGlyLysLeuAlaProAspGlyLysAlaIleThrValHisThrGlnAlaAspThr 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548 GCCTTCGCCGCGGGCTGGCACCTGACTGTGCAATAACCGTAGTACATACCCAGGGACGATACG 607

Qy   141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   608 TCCAGTCCGGATCCGTTATTTAATCCCTCTAAAAACTGGCGTTTGCCAGTGACTAACGCG 667

Qy   161 AsnValThrAspAlaIleLeuIleLeuGluArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668 AACGTGACTGACGCGATCCTCAGCCAGGGCGAGGAGGGTCAATTGCTAGACTTTACGGGCAT 727

Qy   161 AsnValThrAspAlaIleLeuIleLeuGluArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668 AACGTGACTGACGCGATCCTCAGCCAGGGCGAGGAGGGTCAATTGCTAGACTTTACGGGCAT 727

Qy   181 TyrGlnThrAlaPheArgGluLeuGluArgValLeuValLeuAsnPheProGlnSerAsnLeuCys 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   728 CGGCAAACGGGCGTTTCCGGAACTGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC 787

Qy   201 LeuLysArgGluLysGlnAlaAspGluSerCysSerLeuThrGlnValAlaLeuProSerGluLeu 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   788 CTTAAACGTGAGAAACAGGACGAGAGCTGTTCATTAACGCAGGTCGCATTACCATCGGAACTC 847

Qy   221 LysValSerAlaAspCysValSerLeuThrGlyValAlaValSerLeuAlaSerMetLeuThr 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   848 AAGGTGAGCGCGCCCACAATGTCTCATTAACCGGTGCGGTTAAGCTCGATCAATGCTGACG 907

Qy   241 GluIlePheLeuLeuGluGlnInAlaGlnInGlyMetProGlyProGlyTrpGlyArgIleThr 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   908 GAGATATTTCTCCTGCAACAAGCACACAGGGAATCCCGGACGCCTGGGGAAGGGATCACC 967

Qy   261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnInPheAspLeuLeuGln 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   968 GATTCACACCAGTGGAACACCCTTGCTAAGTTTGCATAACGCGCAATTTTGTTTGCTACAA 1027

Qy   281 ArgThrProGluValAlaLeuArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1028 CGCACGCCAGAGGTTGCCCGCAGCCGCACCCCGTTATTAGATTTGATCAAGACAGACG 1087

Qy   301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuValThrSerValLeu 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1088 TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGCGTGACATTACCAGTTCAGTGGCTG 1147

Qy   321 PheIleAlaGlyHisAspThrAsnLeuValLeuAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1148 TTTATCGCCGGACACGATACTAATCTCGGCAAAATCCTCGGAGCTCAACTGG 1207

Qy   341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGluArgTrp 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1208 ACGCTTCCCGGTCAGCCGGATAAACACCCCGCAGGTGGTGAACTGGTGTTTGAACGCTGG 1267

Qy   361 ArgArgLeuSerAspAsnSerGlnInTrpIleGlnInValLeuSerLeuValPheGlnInThrLeuGln 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1268 CGTCGGCTAAGCGGATAACAGCCAGTGGATTCAGGTTTTGTCGCTGGTCTTCCAGACTTTACAG 1327

Qy   381 GlnMetArgAspProLysThrProLeuSerLeuAsnThrProProGlyGlyGluValValLeuSerLeuThr 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1328 CAGATGCGTGATACCGCGCTGTCATTAAATACCGCCGCGGTGAAAGTGACC 1387

Qy   401 LeuAlaGlyCysGluGluArgGlnAsnAlaGlyMetCysSerSerLeuAlaGlyPheThrGln 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1388 CTGGCAGGATGTGAGCAGGCGGAAAATGCCGGCCATGTGTTCGTTGCCAGGTTTTACGGCAA 1447

Qy   421 IleValAlaAsnGluAlaArgIleProAlaCysSerLeu 432
           |||||||||||||||||||||||||||||||||||||
Db 1448 ATCGTGAATGAAGCAGCCATACCGGCGGTGCAGTTTG 1483

RESULT 12
US-10-601-319-7
```
; Sequence 7, Application US/10601319
; Publication No. US20040091968A1
GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O'Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING THEM
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn for Windows Version 4.0
;
SEQ ID NO 7
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (188)...(1483)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 403
; OTHER INFORMATION: n = A,T,C or G
US-10-601-319-7
```

US-10-601-319-10 (1-432) x US-10-601-319-7 (1-1901)

Alignment Scores:
```
Pred. No.:
Score:                  2182.00
Percent Similarity:     98.1%          Length:        1901
Percent Identity:       98.1%          Matches:       424
Best Local Similarity:  96.6%          Conservative:  0
Query Match:            96.6%          Mismatches:    8
DB:                     5.22e-246      Indels:        8
                                       Gaps:          0
```

```
Qy    1 MetLysAlaIleLeuIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
          | | | | | | | | | | | | | | | | | | | |
Db  188 ATGAAAGCGATCTTAATCCTCATTCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAGTCTGCA 247

Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db  248 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCGGTCATCGGT 307

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db  308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 367

Qy   61 ThrTrpProValLysLeuGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db  368 ACCTGGCCGGTAAAACTCGGTGAGCTGACACCGCGCGGTGGTGAGCTAATCGCCTATCTC 427

Qy   81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
Db  428 GGACATTACCAACGCCAGCGTCTGGTGGCCGATGGCCTGCTGCCGAAAAAGGGCTGCCCG 487

Qy  101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db  488 CAGTCTGGTCAGGTCGCCATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA 547

Qy  121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db  548 GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAAATAACCGTTCACATACCCAGGCAGATACG 607

Qy  141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db  608 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACCGGCGTTTGCCAACTGGATAACGCG 667

Qy  161 AsnValThrAspAlaIleLeuGluArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db  668 AACGTGACTGACGCGATCCTCGAGCGGGCAGGCGGAAGCATTGCTGACTTTACCGGGCAT 727

Qy  181 TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db  728 CGGCAAACGCGGTTTCGCGAACTGGAACGGGTGCTTAATTTCCCGCAATCAAACTTGTGC 787

Qy  201 LeuLysArgGlyLysGlnAspGluSerCysSerLeuThrGlnIleAlaLeuProSerGluLeu 220
Db  788 CTTAAACGTGAGAAACGAGGAAAAGCTGTTCATTTAACGCAGATTGCTTTACCATCGGAACTC 847

Qy  221 LysValSerAlaAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db  848 AAGGTGAGCGCCGCAGACTGTGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 907

Qy  241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGlyPheTrpGlyArgIleThr 260
Db  908 GAGATATTTCTGCTGCAACAAGCACAGGGAATGCCGGGATTCTGGGGAAGGATCACC 967

Qy  261 AspSerHisGlnThrTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnProHisLeuLeuGln 280
Db  968 GATTCACCCAGACCTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTATTTGCTACAA 1027

Qy  281 ArgThrProGluValAlaAlaAspThrProLeuProLeuLeuAspIleIleLysThrAla 300
Db 1028 CGCACCCCAGAGGTTGCCGCGGACACCCCGTTATTAGATTTGATTAAAGACAGCG 1087

Qy  301 LeuThrProHisHisProProLysGlnAlaTyrGlyValThrThrLeuProThrSer 320
Db 1088 TTGACGCCCCATCACCCGCCAAAACAGGCGTATGGTGTGACAACTTTACCGACGTCG 1147

Qy  321 PheIleAlaGlyHisAspProThrAsnLeuAlaAsnLeuGlyTyrAlaLeuGlyLeuAsnTrp 340
Db 1148 TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCTACGCGCTGGGGCTCAACTGG 1207

Qy  341 ThrLeuProGlyGlnProAspAlaAsnThrProProGlyLysGlyLeuValAlaArgTrp 360
Db 1208 ACGCTTCCCGGTCAGCCCGGATAACACCCCCGCCAGGTGGCTGAACTGGTGTTTGAACGGCTGG 1267

Qy  361 ArgArgLeuSerAspAlaAsnSerGlnInTrpIleGlnValPheGlnInThrLeuGln 380
Db 1268 CGTCGGCTAAGCGATGCAAACAGCCAGTGGATTCAGGTTTTCGCTCAGACTTTACAG 1327

Qy  381 GlnMetArgGlyLysThrProLeuLeuAspLeuThrProProSerLeuLeuThr 400
Db 1328 CAGATGCGTGGATAAAAACGCCCGTCGTCATTAAAATACGGGGTTTTTAAAAACTGACC 1387

Qy  401 LeuAlaGlyLysTyrCysGlyLeuUArgAsnAlaLeuGlnLeuGlyTyrCysSerSerLeuGln 420
Db 1388 CTGGCAGGATGTGAAGAGCGAAATGCCCAGGGCATGTGTTCGTTCCGTTGGCAGAGG 1447

Qy  421 IleValAsnGlyLeuValLysArgIleProAlaCysSerSerLeu 432
Db 1448 ATCGTGAATGAAGCACGCATTACCGGCGCTGCAGTTTG 1483
```

```
Qy   201 LeuLysArgGlyLysGlnAspGluSerCysSerLeuThrGlnAlaAlaLeuProSerGluLeu 220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   788 CTTAAACGTGAGAAACAGGACGGAAAAGCTGTTCATTAACGGCAGGCCATTAACGGGAACTC 847

Qy   221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerSerMetLeuThr 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   848 AAGGTGAGCGGCCGACAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 907

Qy   241 GluIlePheLeuLeuGlnGlnIleAlaGlnGlyMetProGluProGlyTyrArgIleThr 260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   908 GAGATATTTCTCCTGCTACAACAGCACAGGGAATGCCGGAGCCCGGGTGGGAAGGATCACC 967

Qy   261 AspSerHisGlnThrTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   968 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTATTTGCTACAA 1027

Qy   281 ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleLeuTyrThrAla 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1028 CGCACGCCAGAGGTTGCCGCGAGCCGCCCGTTATTAGATTTGATCAAGACAGCG 1087

Qy   301 LeuThrProHisProProGlyLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1088 TTGACGCCCCATCCACCGGCAAAAACAGGCGTATGGGTGTGACATTACCCACTTCAGTGCTG 1147

Qy   321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGluLeuAsnTrp 340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1148 TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGCCACTGGAGCTCAACTGG 1207

Qy   341 ThrLeuProGlyGlnProAspProAsnThrProGlyGlyLysGlnAlaThrLysValArgTrp 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1208 ACGGCTTCCCGGGTCAGCCGGATAACACGCCGGGAGGTGGGAACCTGGTCGTTTGAAACGGCTGG 1267

Qy   361 ArgArgLeuSerAspAsnSerGlnIleTrpIleGluValSerValPheGlnThrPheLeuGln 380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1268 CGTCGGCTAAGCGATAACAGCCAGTGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1327

Qy   381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProGlyGlyLysValLysLeuThr 400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1328 CAGATGCGTGATAAAACGCCCCTGTCATTAAATACCGGCGGCGGAGGTGAAACGTGACC 1387

Qy   401 LeuAlaGlyTyrGluGlyThrArgAsnLeuAlaGlnGlyMetCysSerSerLeuAlaGlyTyrLeu 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1388 CTGGCAGGATATGGAAAGGCGAAAATGGGCAGGGCATTGTGTTCGTTCGGCAGGTTTAGGCAA 1447

Qy   421 IleLeuValAsnGluAlaArgIleProAlaCysSerSerLeu 432
         ||||||||||||||||||||||||||||||||||||
Db  1448 ATCGTGAATGAAGCAGCCGCATACCGGCCGTGCAGTTTG 1483
```

```
RESULT 13
US-10-933-115-7
; Sequence 7, Application US/10933115
; Publication No. US20050281792A1
; GENERAL INFORMATION:
;   APPLICANT: Short, Jay M.
;   APPLICANT: Kretz, Keith A.
;   APPLICANT: Gray, Kevin A.
;   APPLICANT: Barton, Nelson Robert
;   APPLICANT: Garrett, James B.
;   APPLICANT: O' Donoghue, Eileen
;   APPLICANT: Mathur, Eric J.
;   TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
;   TITLE OF INVENTION: THEREOF
;   FILE REFERENCE: 564462001822
;   CURRENT APPLICATION NUMBER: US/10/933,115
;   CURRENT FILING DATE: 2004-09-01
;   PRIOR APPLICATION NUMBER: US/09/866,379
;   PRIOR FILING DATE: 2001-05-24
;   PRIOR APPLICATION NUMBER: US 09/580,515
;   PRIOR FILING DATE: 2000-05-25
;   PRIOR APPLICATION NUMBER: US 09/318,528
;   PRIOR FILING DATE: 1999-05-25
;   PRIOR APPLICATION NUMBER: US 09/291,931
```

```
;   PRIOR FILING DATE: 1999-04-13
;   PRIOR APPLICATION NUMBER: US 09/259,214
;   PRIOR FILING DATE: 1999-03-01
;   PRIOR APPLICATION NUMBER: US 08/910,798
;   PRIOR FILING DATE: 1997-08-13
;   NUMBER OF SEQ ID NOS: 10
;   SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 7
;   LENGTH: 1901
;   TYPE: DNA
;   ORGANISM: Escherichia coli
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (188)...(1483)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 403
;   OTHER INFORMATION: n = A,T,C or G
US-10-933-115-7

Alignment Scores:
Pred. No.:
Score:                        5.22e-246
Percent Similarity:           2182.00          Length:          1901
Best Local Similarity:          98.1%          Matches:          424
Query Match:                    98.1%          Conservative:       0
DB:                             96.6%          Mismatches:         8
                                               Indels:             0
                                               Gaps:               0

US-10-601-319-10 (1-432) x US-10-933-115-7 (1-1901)
```

```
Qy     1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTTCCTGATCCCGTTAAACCCGCAATCTGCA 247

Qy    21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248 TTCGCTCAGAGTGAGCCCGGAGCTGAAAGTGTCGTGATTGTCAGTCGTCATGGT 307

Qy    41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGACGTAACGCCAGATGGGCCA 367

Qy    61 ThrTrpProValProLeuGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368 ACCTGGCCGGTAAAACTCGGTGAACTGACACCCGCNGGTGGTGAGCTAATCGCCTATCTC 427

Qy    81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 GGACATTACCAACGCCAGCGGTCTGGTAGCCGACGGATTGCTGCCGAAAAGGGCTGCCCG 487

Qy   101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488 CAGTCTCGGTCAGGTCGCGATAATTGCCGATGTCGACGAGCGTACCCGTAAAAACAGGCGAA 547

Qy   121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548 GCCTTCGCGCGCCGGGCTGGCACCTGACGTGCAATAACGGCACCATCCAGGCAGATACG 607

Qy   141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   608 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG 667

Qy   161 AsnValThrAspAlaIleLeuGluGlyArgAlaGlyGlySerIleGluAlaAspProPheThrGlyHis 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668 AACGTGACTGACGCGATCCTTGAGGGCAGGGCAGGGGTCAATTGCCTGACTTTACCGGGCAT 727

Qy   181 TyrGlnThrAlaPheAlaArgGluLeuAsnGlnValLeuAsnProHeSerAsnLeuCys 200
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   728 CGGCAAACGGCCGTTTCCGGAACTGAACGGGGTTAATTTCCGGCAATCAAACTTTGTGC 787

Qy   201 LeuLysArgGlyLysGlnAspGluSerCysSerLeuThrGlnAlaAlaLeuProSerGluLeu 220
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   788 CTTAAACGTGAGAAACAGGACGGAAAAGCTGTTCATTAACGGCAGGCCATTAACGGGAACTC 847
```

```
;   PRIOR FILING DATE: 1999-03-01
;   PRIOR APPLICATION NUMBER: US 08/910,798
;   PRIOR FILING DATE: 1997-08-13
;   NUMBER OF SEQ ID NOS: 4
;   SOFTWARE: Patentin version 3.1
;
;   SEQ ID NO 3
;   LENGTH: 1901
;   TYPE: DNA
;   ORGANISM: Escherichia coli appA phytase
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (403)..(403)
;   OTHER INFORMATION: n is any nucleotide
US-11-056-354-3

US-10-601-319-10 (1-432) x US-11-056-354-3 (1-1901)

Alignment Scores:
Pred. No.:
Score:                    5.22e-246
                          2182.00          Length:          1901
Percent Similarity:       98.1%            Matches:         424
                          98.1%            Conservative:    0
Best Local Similarity:    98.1%            Mismatches:      8
Query Match:              96.6%            Indels:          0
DB:                       14               Gaps:            0
```

```
Qy    1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  188 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 247

Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  248 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 307

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  308 GTGCGTGCGCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 367

Qy   61 ThrTrpProValLysLeuGlyLeuGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  368 ACCTGGCCGGTAAAACTGGGTTTGGGAGCTGACACCGCGNGGCGGTGAGCTTAATCGCCTACTC 427

Qy   81 GlyHisTyrTrpAlaArgGlnArgLeuValAlaLeuAspGlyLeuLeuProLysCysGlyCysPro 100
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  428 GGACATTACCAAACGCCACGCGGTCTGGGTAGCCGACGGATTGCTGGGCGAAAAAGGCTGCCG 487

Qy  101 GlnSerGlyGlnValAlaIleIleAlaAspValAspArgThrArgGlyThrGlyGlu 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  488 CAGTCTGGTCAGGTCGGCGGATTATTGCTGATGTGCGACCCGTAACCAACAGGCGGAA 547

Qy  121 AlaPheAlaAlaGlyTyrLeuAlaProArgAspAspCysAlaIleThrValHisThrGlnAlaAspThr 140
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  548 GCCTTCGCCGCCGGTGGCCACCTGACCGTGCAAATAACCGTACATACCAGCAGATACG 607

Qy  141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysSerLeuAspAspAsnAla 160
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  608 TCCAGGTCCCGGATCCGGTTATTTAATCCCTCTAAAAAACNCGGCGGTTTACGGTAGCGCG 667

Qy  161 AsnValThrAspAlaIleLeuGluAsnAlaGlyAlaGlySerIleLeuAlaAspProPheThrThrGlyHis 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  668 AACGTGACTGACCGGCATCCTCAGCAGGGCAGGAGGGGTCAATTGCTGACTTTACCGGGCCAT 727

Qy  181 TyrGlnThrAlaPheAlaGlnLeuGluLysValLeuAsnPheProGlnSerAsnLeuCys 200
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  728 CGGCAAACGGGCGTTTCCCGAACTGGGCTTAATTCAAAACTTCTGC 787

Qy  201 LeuLysArgGlnLysAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  788 CTTAAACGTGAGAAACAGGAACACGAAAGCGTGTTCATTAACGGCAATTACCATCGGAACTC 847

Qy  221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  848 AAGGTGAGCGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCTCGCATCAATGCTGACG 907
```

```
Qy  221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  848 AAGGTGAGCGCGCCGACAATGTCTCATTAACCGGTGCGGTAAGCTCGCATCAATGCTGACG 907

Qy  241 GluIleProheLeuLeuGlnGlnGlnnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  908 GAGATATTTCCTTCCTGCAACAAGCAACACAGGGAATGCCGGAGCCGGGTTGGGGAGGATCACC 967

Qy  261 AspSerHisGlnTrppAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  968 GATTCACCCAGTGGAACACCTTGCTAAGTTTGCATAACGGGCAATTTATTTGCTACAA 1027

Qy  281 ArgThrProGluValAlaLaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1028 CGCACGCCAGAGGTTGCCCGCCCAGCCCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 1087

Qy  301 LeuThrProHisProProGlnLysGlnValThrGlyValThrLeuProThrSerValLeu 320
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1088 TTGACGCCCCATCCACCGCAAAAACAGGGCGTATGGGTGACATTACCGACTTCAGTGCTG 1147

Qy  321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1148 TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGAGCTCAACTGG 1207

Qy  341 ThrLeuProGlyGlnProAspAspAsnThrProGlyGlyLeuValPheGluLeuArgTrp 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1208 ACGCTTCCCGGTCAGCCCGGATAACACGCCCGGCGGTTGGTGTTGAACGGCTGG 1267

Qy  361 ArgArgLeuSerArgAspAsnSerGlnIleThrIleGluGlnValSerLeuValPheGlnLeuGln 380
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1268 CGTCGGCTAAGCGATAACAGCCAGATCACGATTGAGCAGGTTTCGCTTGTCTTTCAGCTTCAG 1327

Qy  381 GlnMetArgAspProLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuLeuThr 400
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1328 CAGATGCGTGATAAAACCCCGCTGTCATTAAATACCCCGCCAGGCGAAGTGAAACTGTTGACC 1387

Qy  401 LeuAlaGlyCysGluGluGlnArgAsnAlaGlnGlyMetCysSerLeuValAlaGlyPheThrGln 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1388 CTGGCAGGATGTGAAGAGCAAAATGCGCAGGGCATGTGTTCGTTGGTCGCAGGTTTTACGCAA 1447

Qy  421 IleValAsnGlyIleAlaArgIleLeuProAlaCysSerSerLeu 432
        |||||||||||||||||||||||||||||||
Db 1448 ATCGTGAATGGAATGCAACGCCATACCGGCGGTGCAGTTTG 1483
```

```
Qy    241  GluIlePheLeuLeuGlnGlnIleAlaGlnGlyMetProGlnProGlyTrpGlyArgIleThr  260
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908  GAGATATTTCTCCTGCAACAACGCACAGGGAATGCCGCAGCCCGGGGGGAAGGATCACC  967

Qy    261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln  280
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968  GATTCACACCAGTGGAACACCTTGCTTAAGTTGCATAACGCGCAATTTTATTTGCTACAA  1027

Qy    281  ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuLeuLysThrAla  300
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1028 CGCACGCCAGAGGTTGCCCGCAGCGCCGACCCCGTTATTAGATTTGATCAAGACAGCG  1087

Qy    301  LeuThrProHisProProGlnLysGlnAlaThrGlyValThrLeuProThrSerValLeu  320
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1088 TTGACGCCCCATCCACCGCAAAAACAGGGTATGGTGTGGTGACATTACCCACTTCAGTGCTG  1147

Qy    321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyLysAlaLeuGluLeuAsnTrp  340
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1148 TTTATCGCGGACACGAATCTAATCTGCGAAAATCCGGCGGGGAAGGCTCAACTGG  1207

Qy    341  ThrLeuProGlyGlnProAspAspAsnThrProGlyGlyLeuGluValValAlaIleArgHisGly  360
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1208 ACGGTTCCCCGTCAGCCCGGATAACACACGCCGGGGGGGCTGGAAGTGGTCGCTCAAGGT  1267

Qy    361  ArgArgLeuSerAspAsnAsnSerGlnInTrpIleGlnInValSerLeuValPheGlnInThrLeuGln  380
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1268 CGTCGGCCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTCGGTGTCTTCCAGACTTTACAG  1327

Qy    381  GlnMetArgAspGlyLysThrProLeuSerLeuAsnThrProProGlyGluValGlyGlu  400
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1328 CAGATGCGTGATGGTAAAAACCCGCTGTCATTAAATACGCCCCCGGGAGAGGTGAAACTGACC  1387

Qy    401  LeuAlaGlyCysGluGlnValArgGlnAsnLeuGlnInGlyMetCysSerSerLeuAlaGlyTyrPheThrGln  420
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1388 CTGGCAGGAATGTGAAGCAGCGAAATGCCAGCCAGGCCATAGTGTCGTTCGTTGGCAGGTTTTACGCAA  1447

Qy    421  IleValAsnGluValAlaArgIleProAlaCysSerLeu  432
                ||||||||||||||||||||||||||||||||||||||||
Db    1448 ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTG  1483
```

```
Alignment Scores:
Pred. No.:
Score:                  6.85e-246
                        2181.00       Length:        1901
Percent Similarity:     98.1%         Matches:        424
Best Local Similarity:  98.1%         Conservative:    0
Query Match:            96.6%         Mismatches:      8
DB:                     3             Indels:          0
                                      Gaps:            0
```

US-10-601-319-10 (1-432) x US-09-866-379-5 (1-1901)

```
Qy      1  MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla  20
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188  ATGAAAGCGGATCTTAATCCCATTTTAATTCCTTCTGATTCCGTTAACCCCGCAATCTGCA  247

Qy     21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly  40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248  TTCGCTCAGAGTGAGCCGGAGCTGAAAACTGGGGGTGTCAGTCGTCGTCATGGT  307

Qy     41  ValArgAlaAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro  60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308  GTGCGTGCTCCAACCAAGGCCACCCAACTGATGCAAGATGTCACCCCAGACGCATGGCCA  367

Qy     61  ThrTrpProValLysLeuGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu  80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    368  ACCTGGCCGGTAAAACTGGGTGAACTGACACCCGCGTGAGTCGAATCGGCCTATTCTC  427

Qy     81  GlyHisLeuTyrArgGlnAsnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro  100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428  GGACATTACCAACGCCAGCCTCGGGTGAGCCGACGGATTGCTGGGCGAAAAGGGGCTGCCCCG  487

Qy    101  GlnSerGlyGlnValAlaIleIleIleAlaAsnAspValArgThrArgGlyThrGlyGlu  120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488  CAGTCTGGGTCAGGTCGCGATTATTGCTAGTGCACGACGTACCGTAAAAACAGGCGAA  547

Qy    121  AlaPheAlaAlaGlyLeuAlaProAspAspAspCysAlaIleThrValHisGlnAlaAspThr  140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548  GCCTTCGCCGCCGGCCTGGCACCTGACATTGTGACCGTTCCGGCAGCGGCAGATACG  607

Qy    141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAlaAsnAla  160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608  TCCAGTCCCGATCCGTTATTTAATCCCTCTAAAAAACTGGGGGTTTGCCAACTGGATAACGCG  667

Qy    161  AsnValThrAspAlaIleLeuGluArgAlaGlyGlySerIleAlaAspPheThrGlyHis  180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    668  AACGTGACTGACGCCGGATCCTCAGCAGCGGCGAGGGGTCAAATTGCTGACTTTACCGGCAT  727

Qy    181  TyrGlnThrAlaPheArgGluLeuGluLeuValLeuAsnPheProGlnSerAsnLeuCys  200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    728  CGGCAACGACGCGTTTCGCGAACTGGAACAGGGTGCTTAATTTCCGCAATCAAACTTGTGC  787

Qy    201  LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    788  CTTAAAACGTGAGAAAACAGGACGAAAGCTGTTCATTACCAGGCATTACCATCGGAACTC  847

Qy    221  LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr  240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    848  AAGGTGACCGCCCATCATTAACCGGTGCGGTAAGCCTCGCCATCAATGCTGACG  907

Qy    241  GluIlePheLeuLeuGlnGlnIleAlaGlnGlyMetProGlnProGlyTrpGlyArgIleThr  260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908  GAGATATTTCTCCTGCAACAACGCACAGGGAATGCCGCAGCCCGGGGGGAAGGATCACC  967

Qy    261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln  280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968  GATTCACACCAGTGGAACACCTTGCTTAAGTTGCATAACGCGCAATTTTATTTGCTACAA  1027

Qy    281  ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuLeuLysThrAla  300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028  CGCACGCCAGAGGTTGCCCGCAGCGCCGACCCCGTTATTAGATTTGATCAAGACAGCG  1087
```

```
Db   1028  CGCACGCCCAGAGGTTGCCCCCAGCCGCACCCCGTTATTAGATTTGATCAAAGACAGCG 1087

Qy    301  LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu  320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088  TTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCTG 1147

Qy    321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaAlaLeuGluLeuAsnTrp  340
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148  TTTATCGCCCGGACACCGATACTAAATTCTCGGCAAATCTCCGGCGGCCACTGGGAGCTCAACTGG 1207

Qy    341  ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp  360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208  ACGCTTCCCGGTCAGCCGGATAACACGCCGCCGGATGGTGAAACTGGTGTTTGAACGCTGG 1267

Qy    361  ArgArgLeuSerAspAsnSerGlnnTrpIleGlnValSerLeuValPheGlnThrLeuGln  380
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268  CGTCGGCTAACGGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1327

Qy    381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr  400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1328  CAGATGCGTGATAAAAACGCCCGCTGTCATTAAATACCGCCCGGAGAGGTGAAACTGACC 1387

Qy    401  LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerSerLeuAlaGlyPheThrGln  420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388  CTGGCAGGGATGTGAAGAGCGAAAATGGGCAGGGCATGTGTTCGTTGGCAGGTTTTACGGCAA 1447

Qy    421  IleValAsnGluAlaArgIleProAlaCysSerLeu  432
            ||||||||||||||||||||||||||||||||||
Db   1448  ATCGTGAATGAAGCACGCCATACCGGCGGGCTGCAGTTTG 1483
```

Search completed: June 14, 2006, 15:10:36
Job time : 1441.42 secs

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

Run on: June 14, 2006, 14:22:06 ; Search time 33.1927 Seconds
                                         (without alignments)
                                         4595.351 Million cell updates/sec

Title: US-10-601-319-10
Perfect score: 2258
Sequence: 2258
Sequence: 1 MKAILIPFLSLLIPLTPQSA..........CSLAGPTQIVNEARIPACSL 432

Scoring table:
                BLOSUM62
                Xgapop   10.0 ; Xgapext   0.5
                Ygapop   10.0 ; Ygapext   0.5
                Pgapop   6.0 ; Pgapext   7.0
                Delop   6.0 ; Delext   7.0

Searched: 290227 seqs, 117694381 residues

Total number of hits satisfying chosen parameters: 580454

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10601319/runat_12062006_180105_27789/app_query.fasta_1
-DB=Published_Applications NA_New -QFMT=fastap -SUFFIX=rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abss03h
-USER=US10601319_@CGN_1_28_@runat_12062006_180105_27789 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :   Published_Applications NA_New:*
            1: /EMC_Celerra_SIDS37/pcodata/2/pubpna/US09_NEW_PUB.seq:*
            2: /EMC_Celerra_SIDS3/pcodata/2/pubpna/US06_NEW_PUB.seq:*
            3: /EMC_Celerra_SIDS3/pcodata/2/pubpna/US06_NEW_PUB.seq:*
            4: /EMC_Celerra_SIDS3/pcodata/2/pubpna/US07_NEW_PUB.seq:*
            5: /EMC_Celerra_SIDS3/pcodata/2/pubpna/US08_NEW_PUB.seq:*
            6: /EMC_Celerra_SIDS3/pcodata/2/pubpna/PCT_NEW_PUB.seq:*
            6: /EMC_Celerra_SIDS3/pcodata/2/pubpna/US10_NEW_PUB.seq:*
            7: /EMC_Celerra_SIDS3/pcodata/2/pubpna/US11_NEW_PUB.seq:*
            8: /EMC_Celerra_SIDS3/pcodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	%	Description
1	111	4.9	2661	6	US-10-505-928-848	Sequence 848, App
2	110	4.9	1482	6	US-10-449-902-14217	Sequence 14217, A
3	107	4.7	1519	6	US-10-449-902-9793	Sequence 9793, Ap
4	106.5	4.7	1470	6	US-10-449-902-21879	Sequence 21879, A
5	106.5	4.7	2822	6	US-10-449-902-7723	Sequence 7723, Ap
6	106.5	4.7	118899	7	US-11-189-279-64	Sequence 64, Appl
7	105	4.7	1221	7	US-11-284-978-14	Sequence 14, Appl
8	105	4.7	271	7	US-11-145-307A-189	Sequence 189, App
9	104.5	4.6	12225	6	US-10-501-834-1	Sequence 1, Appl

10	101.5	4.5	3646	6	US-10-511-937-358	Sequence 358, App
11	99	4.4	39536	7	US-11-293-697-315	Sequence 315, Appl
12	98.5	4.4	2696	7	US-10-857-260-27	Sequence 27, Appl
13	96.5	4.3	2696	7	US-11-317-329-2	Sequence 2, Appl
14	96.5	4.3	2696	7	US-11-317-329-3	Sequence 4, Appli
15	96.5	4.3	2696	7	US-11-317-329-5	Sequence 5, Appli
16	96.5	4.3	2696	7	US-11-317-329-7	Sequence 7, Appli
17	96	4.3	1007	6	US-10-449-902-9589	Sequence 9589, Ap
18	96	4.3	4239	6	US-10-511-937-543	Sequence 543, App
19	96	4.3	274419	6	US-10-857-260-34	Sequence 34, Appl
20	95.5	4.2	2696	7	US-11-317-329-8	Sequence 8, Appli
21	95.5	4.2	5193	6	US-11-270-040-5	Sequence 5, Appli
22	94.5	4.2	5449	6	US-10-196-749-351	Sequence 351, App
23	94.5	4.2	4407	6	US-10-505-928-800	Sequence 800, App
24	94	4.2	1945	7	US-11-293-697-351	Sequence 351, App
25	93.5	4.1	1945	7	US-11-293-697-1957	Sequence 1957, Ap
26	93.5	4.1	2019	7	US-11-293-697-1629	Sequence 1629, Ap
27	93.5	4.1	3086	7	US-10-449-902-12631	Sequence 12631, A
28	93	4.1	2622	7	US-11-293-697-95	Sequence 95, Appl
29	93	4.1	3731	6	US-10-449-902-27838	Sequence 27838, A
30	92.5	4.1	1314	6	US-10-449-902-3379	Sequence 3379, Ap
31	92.5	4.1	1801	6	US-10-449-902-15401	Sequence 15401, A
32	92.5	4.1	2696	7	US-11-317-329-9	Sequence 9, Appli
33	92.5	4.1	2696	7	US-11-317-329-10	Sequence 10, Appli
34	92	4.1	1488	6	US-10-449-902-24588	Sequence 24588, A
35	92	4.1	2422	6	US-10-449-902-24390	Sequence 24390, A
36	92	4.1	3623	6	US-10-449-902-9933	Sequence 9933, Ap
37	91.5	4.1	1870	6	US-10-449-902-20180	Sequence 20180, A
38	91.5	4.1	3107	7	US-11-293-697-215	Sequence 215, App
39	91.5	4.1	4094	6	US-10-449-902-13890	Sequence 13890, A
40	91.5	4.1	6400	6	US-10-449-902-396	Sequence 396, App
41	91	4.0	1903	6	US-10-449-902-23156	Sequence 23156, A
42	91	4.0	1987	7	US-11-293-697-1200	Sequence 1200, Ap
43	91	4.0	2359	6	US-10-449-902-8080	Sequence 8080, Ap
44	90.5	4.0	1526	6	US-10-449-902-13489	Sequence 13489, A
45	90.5	4.0	1737	6	US-10-449-902-19336	Sequence 19336, A

ALIGNMENTS

RESULT 1
US-10-505-928-848
; Sequence 848, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: Patentin 3.2
; SEQ ID NO 848
; LENGTH: 2661
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-505-928-848

Alignment Scores:
Pred. No.:   0.354
Score: 111.00   Length: 2661
Percent Similarity: 35.6%   Matches: 115
Best Local Similarity: 22.2%   Conservative: 69
Query Match: 4.9%   Mismatches: 174
DB: 6   Indels: 160
                                                      Gaps: 30

US-10-601-319-10 (1-432) x US-10-505-928-848 (1-2661)

Qy   40 GlyValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAsp------ 57
         ||:::| | ||||| || |||| ||| |::||| ::|:| ||| || || |||
Db   253 GGACTGCCAGCTCCCTGGGGTCTCCTTCTTCTGGGGGCAGAATCCTCAGTCCTCCTTGACT 312

```
Qy         58 ------------------ThrTrpProValLysLeuGly------------------------------ 67
Qy        313 TCACGACTGTGGCCAGATCANTGTGTGACACTGTTTGGGGTGTTCCAGAGCGCTTG                     372
Db         68 GluLeuThrProArgGlyGlyValGlyLeuLeuAlaTyrLeuGlyHisTyrTrpArgGlnArg               87
Db        373 CATCAAACACCCTAACTCAGAAGTGTGCAGCCACACTGGGACTCGAAACCCAACAACAG                   432
Qy         88 ------------------------LeuValAlaAlaAspGlyLeuLeu---Pro 95
Db        433 GGACGAGAGAGACTCACGCCCTTGGGGTGCCCGGGTTCTCGGTGCAATCAGGCATGACTTCCAGC            492
Qy         96 LysCysGlyCysSerProGlnSerGlyGlyValValAlaIleIleLeuAlaAspGluArg               114
Db        493 TCCTGCGCCTT-CCCCAGCAA---CTGCTGACTGGGGACCCCAGGAGGGAGCTG                        542
Qy        115 ThrArgLysThrGlyGluIleAlaPheAlaAlaAlaGlyLeuAlaProAlaAspCysAlaIleThrVal        134
Db        543 AGCGACGGGGCCTGGGCTGGAGCGGAAGCTCGGGGTCT                                        575
Qy        135 HisThrGlnAlaAspThrSerProAspPro-----LeuPheAsnProLeuLysThr 152
Db        576 CACTCAGGCACCACAGCCCCTCCTTGGCCCCAGGGCTTGAGTGACTCACAGCCTGATTTCAGGCA            635
Qy        153 GlyValCysGlnLeu------------------AspAsnAlaAsnValThr------ 163
Db        636 GGAGCTGCTCTTCTTCTGGGGTATCGCGATCCACTTAAGGACATGAGGCAGACTTCGTTGACAAGC          695
Qy        164 ------------------------------------AspAlaIleLeuGluArg 169
Db        696 TGGTCTGACGACGGCTTCCAGAGCCAGAACTGAGCCCCAGTGAGAGGAGCGACCACCTGGAGCAG           755
Qy        170 AlaGlyGlySerIleAlaAlaAspPheThrGly----------------------His 180
Db        756 CCTGGA-----------TTCCTGGGTGTCCCCGGCAGCCACACACAGCCATGCAC                        800
Qy        181 TyrGlnThrAlaPhe--------------------ArgGluLeuGlnArgValLeu 192
Db        801 TACCCAACTGCACTCTCTTCCTCCATCCTGGCCAATGGGGCCCCAGGCCTTTCGGCAATCTGC            860
Qy        193 AsnPheProGlnSerAsnLeuCysLeuLysArgGlnLysGlnAspGluSerCysSerLeu 212
Db        861 GCCTTCAAATGCCCAGGCGGCCTGACACTGGCCAAGGTGGCCCAGGGAG                            905
Qy        213 ThrGlnAlaLeuProSerGluLeuLeuLysValAlaSerAlaAspCys----------- 226
Db        906 --CAGGTGATGGAACACCTTAGTTCGGGATAACTNGGCTCGGTGACATCANTGGGCTGCAG              962
Qy        227 ----------ValSerLeuThrGlyAlaValSerGluAlaSerMetLeuThrGluIlePheLeu 244
Db        963 GAGGTGGTGAGACTCTTCCGGCAGG-----GCCANCCCCTCCTTCGGCTTCGACAGA-              1010
Qy        245 LeuGlnIleGlnAlaGlnGly------------------MetProGluProGlyTrpGlyArg 258
Db       1011 CTCAATCGATTTGATNGGCTCTNGGGCCCTACAGCACCCTGAGCAGCCCCCAGTGCGGGGCGC          1070
Qy        259 IleThr-----------------------------AspSerHisGlnThrProAsnThrLeuLeu 269
Db       1071 AGCACCTNACATGGAGACGTATGATGTACTNCGGTCACACAAAACCAGGTCCTGAGT                    1130
Qy        270 Ser---LeuHisAsnAlaGlnPheAspLeuLeuGlnArgThrProGluValAlaArg-- 287
Db       1131 TCCTACGGTGTACAACGATGAGGATGACGTCTTTTGCCGGGAGCCANTTGTGGCCCAGTTC               1190
Qy        288 -------SerArgAlaThrProLeuLeuAspLeuIleLysThrArgLeuThrProHis 304
Db       1191 TCTTTGCCCAGGCAANTGCCTTCCAGCCTGCTGTGCGCTGCACACCACTCCTAAG                      1250
Qy        305 ProProGlnLysGln------------------AlaTyrGlyValThrLeuProThrSer---- 318
Db       1251 GCCGTAGAGAGGAGCTGAACCGACCCTCTACGANGTGTTTTCTGGAGGTCTCCAGCACTGG              1310
Qy        319 ------------------ValLeuPheIleAlaAlaGlyHisAspThrArgAsnLeuValAlaAsnLeuGly--- 333
Db       1311 CAGAGCAAGGAGGACGAGTGATCCTGCTTGGGGGACTTCAANTCTGACTGACGCGCTTCACTGACCAAA          1370
Qy        334 ------------GlyAlaLeuGluLeu----------------AsnTrpThrLeuProGly 344
Db       1371 AAGCGGCCTGGACAAGCTGGAGCTGCGCGACTGAGCCAGGCTTCCACTNGGGTGATTTGCCGAT             1430
Qy        345 GlnProAspAsnThrProProGlyGlu-------------------------LeuVal 356
Db       1431 GGGGGAGGAGGACCACACAGTACGGCGGCCACCCCACTGACATGCCCATTGACCAGGTCTGCTGGTGC        1490
Qy        357 PheGluArgTrpArgArgLeuSerArgSerAsnSerGlnArgProIleGluGlnValValSerLeuValPhe 376
Db       1491 GGGGGAGCGCGGCTGGCGGGAGTCTGCTCGCACACTGCG---------GCTGCCTTGACTTC                1538
Qy        377 GlnThrThrLeuGlnGluMetArgAspLysThrProLeuSerLeuAsnThrPro---ProGly 395
Db       1539 CCCACAGAGCTTCCAGCTCACCGAGGAGGAGGCCCTCAACACATCAGTGACCACTACCCGTG               1598
Qy        396 GluValLysLeuThrLeuAla--GlyCysGluGluArgAsnAlaGlnGlnMetCysSerIle 415
Db       1599 GAGGTGGAGCTGAAGCTAGAGGCCACAGGGCCACCAGCGTCCTCAGCCTCTCAGCCTCACTGTTCTG          1658
Qy        415 euAlaGlyPheThrGlnIleValAsnGluAlaAlaArgIleProAlaCys 430
Db       1659 TTGCTGCTATCACTCCTGTC-----CCCTCAGCTCGTGTGCCCTGCTGC 1699
```

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:

US-10-601-319-10 (1-432) x US-10-449-902-14217 (1-1482)

Length:	1482
Matches:	66
Conservative:	33
Mismatches:	119
Indels:	82
Gaps:	11

Pred. No.: 0.191
Score: 110.00
Percent Similarity: 33.0%
Best Local Similarity: 22.0%
Query Match: 4.9%
Query Match: 6

```
Qy         19 SerAlaPheAlaGlnSerGluProGluLeuLysGluSerValValIleLeuValSerArg 38
Db        957 TCGATCTTCACCTTCTCCGGCGCCCACGGTCAGCTTCCCTTCGGTGGTGAGACCGGGAAC 898
Qy         39 HisGlyValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThr-ProAspAl 58
Db        897 AGCTCGTCCTCGGGTGCGTCACGCACACGCCGGTTGATGTCGGCCAGATCTCCCCT----- 843
```

; SEQ ID NO 9793
; SOFTWARE: Patentin Ver. 2.1
; NUMBER OF SEQ ID NOS: 56791
; PRIOR FILING DATE: 2002-12-11
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-29
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; FILE REFERENCE: MOA-A0205Y1-US
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; APPLICANT: Foundation for Advancement of International Science.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: National Institute of Agrobiological Sciences.
; GENERAL INFORMATION:
; Publication No. US20060123505A1
; Sequence 9793, Application US/10449902

RESULT 3

US-10-449-902-9793/c

; DB: US-10-449-902-9793
; DATABASE ENTRY DATE: 2002-08-28
; DATABASE ACCESSION NUMBER: AK107141
; PUBLICATION INFORMATION:
; ORGANISM: Oryza sativa
; TYPE: DNA
; LENGTH: 1519

Alignment Scores:
Pred. No.:   0.38
Score: 107.00
Percent Similarity: 33.4%   Length: 1519
Best Local Similarity: 21.6%   Matches: 84
Query Match: 4.7%   Conservative: 46
DB: 6   Mismatches: 111
                                 Indels: 148
                                 Gaps: 19

US-10-601-319-10 (1-432) x US-10-449-902-9793 (1-1519)

Qy   58 aTrpProThrTrpProValLysLeuGly------------------------------------GluLe 69
         |||:
Db   842 -TrpGGCGACATGGGAGCAGCGTCTGGATGATGAGGGTCGGAGGGGGACGAGGAGACGAAC 784

Qy   69 uThrProArgGlyTyrGly-GluLeuIleAlaTyrLeuGlyHisTyrTrpArgGlnArgLeuV 89
         ||||||| :::
Db   783 TrGGAGTAGAGGAGGACTCCACCTTGTCCACCGGCCTTCGGACACGAACAGCGAGTAGACGAGG 724

Qy   89 alAlaAspGlyLeuLeuProLysCysGlyCysProGlnSerGlyGln------------------- 104
         |||:||
Db   723 TCCCAGAGATGGACTG----GGAGTCCTTCACGGTGGGGGATGCCGTTCACCTCC 676

Qy   105 --ValAlaIleIleAlaIleA 110
                                                     ||:|:::||||||||:||:
Db   675 AGCGTTCGCCTTCGGTCGGGGATGAAGGGGCGGGGAGGGGGATGAAGTAGGACGTTGCCCTTCTTGCCG 616

Qy   110 spValAspGluArgThrArgLysThrGly-----GluLeuPheAlaAlaAlaGlyLeuAla- 127
         |||:   ::: ||||||||||||||
Db   615 ACGTGACGACGGTGTACTCGAGCCGAGCTGCTTGAGCCTCGATGCGGGTTCGGCC 556

Qy   128 -----ProAspCysAlaIleThrValHisThrGlnAlaAlaAspThrSerSerProAsp--- 144
                   |||||
Db   555 TTCTTGAGGACGTTGTTGTTGAGACTGCCGACGAGGCCCGCGCTCCGGTGAGGACGACG 496

Qy   145 -----ProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAsp-----AsnAlaA 161
                   ||||
Db   495 AGGGGCGACCTTCTTGACGGGGCGGATGCGGGGTGAGGGGTCGGATGTCTTCCGGCTGG 436

Qy   161 snValThrAspAlaIleLeuGlnArgAlaGlyGlySerIleAlaAlaSerPheThrGly--- 179
         |||:|| :::
Db   435 ATCTCCTCGGTTCATGTTGTAGAGCGACACTTCCACCAGCGGCCTTCCGAAGAGGGGGGGGAGGAGGAC 376

Qy   180 ------------------------HisTyrGlnThrAlaPheArg 187
                                 :::
Db   375 ACCACCGCCTCCTGGCGCCCGCCACCTTGGCCCACCACCCATCGCCTCCGTG 316

Qy   187 luLeuGlyArgValLeuAsnPheProGlnSerAsnLeuCysLeuLysArgGluLysGlnA 207
         |||
Db   315 ATCTTCTGCGTGTTGCCCAC 289

Qy   207 spGluSerCysSerLeuThrGlnAlaLeuProSerGluLeuLysValSerAlaAspCysV 227
         ||||||||| :::
Db   288 ATCGGCGGTTCCGAGCTACGACGAGGACCCGGATGCGGGGGGGGCCGTGCTGCTGCCG 229

Qy   227 alSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThrGluIlePheLeuLeuGlnG 247
         ||
Db   228 CAGCTAGGAGCGGAGCGGCGGGGCG- 205

Qy   247 lnAlaGlnGlyMetProGlnProGlyTyrTrpGlyArgIleThrAspSerHisGlnTrp 265
         ||| :::
Db   204 GTGGCGCAGGGGCGGGGAGGAGGACCAGGCCGGTGGAGAGGTTCGGAGCGACGACATGGCCTGG 149

Qy   146 LeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAlaAsnValThrAspAla 165
         ||| :::
Db   579 CTCGAGGACGAGAACGG- 556

Qy   166 IleLeu-------GluArgAlaGlyGly------------------------------SerIle 174
         ||| ::: |||:::
Db   555 GTTCTCCATGAGGAGGAGAGCAGGTCGGCGGCCCCGGGGGACAAGCGTGAGGCCCTTCCAGCGTC 496

Qy   175 AlaAsp-PheThrGlyHisTyrGlnThrAlaPheArgGluLeuGlyArgValLeuAsnPh 194
         |||:|:   :::
Db   495 GCAGAGCCGCGTCGGCTCCTTGGCGTGCCACTTTAAGACGAGACAACTTCGAGGAAGCAGCC 436

Qy   194 eProGlnSerAsnLeuCysLeuLys-----ArgGluLysGlnAspGluSerCysSerLeu-- 212
         ||| :::
Db   435 ATCGAGCAGGAGCATCACGATGAAGTCGTCGAGGCGAGGCGCGCTTCGGCTTGTA 376

Qy   213 ----ThrGlnAlaLeuProSerGluLeuLysValSerAlaAspCysVal---SerLeuTh 230
             ||| :::||||
Db   375 GCAGGCGCGGCGCGCCTCGGGCCTTCCAGGGTTCTCCATCTTCTCGGCGACCAAAGTTGGACGCGGCTGAC 316

US-10-601-319-10 (1-432) x US-10-449-902-9793 (1-1519)

Qy   15 LeuThrProGlnSerAlaPheAlaGlnSerGlnProGluLeuPheLeuGluSerValVal 34
         ||| :::
Db 1065 CTGGCTCATCAGTGCCACGTAGCTCACCAGGAGCCC---CTCCTCCTCGCGGGGGCT 1009

Qy   35 IleValSerArgHisGlyValArgAlaProThrLysAlaThrGlnIleLeuMetGlnAspVal 54
         |||||| :::
Db 1008 CTGCTCGAACGCGAGCAGGTTCGGCGACGACGGGCCGCCCTCGGGTCGTCGATCAGGACTGT 949

Qy   55 --ThrProAspAla 58
                                                           :::
Db   948 GGGTATCCCATTACGGCCCCAGCGGCGGTCGGAAGCCCAGTGCGGTCGTACCTCGGTCCCCCTGC 889

Qy   59 TrpProThrTrpProValLysLeuGlyGlyGlyGlyGluLeuIleAla 78
         ||| :::
Db   888 GGCCGGCCGGTGGGCTTCTGCGCGGCGGACGAACTGTACCCGCCGC---CTCTCT 841

Qy   79 TyrLeuGly------------------------------HisTyrTrpArg 85
         ||| :::
Db   840 CAtCTCGGTTGCGCATGGTATCACCCGGTGGGGTCGGTGCGGTCGGTCGGAGACCG 781

Qy   86 ---Gln 86
                                                                    :::
Db   780 TGGGGCCTCGGCGGCTTCGCGGCGGCGTGTTCGGGTGGGGAATGAAGCTTCTGGTAGTACAGATG 721

Qy   87 ArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysProGlnSerGlyGlnValAla 106
         ||| :::
Db   720 GAGGAGGTGGTGGAC----CTCCCAGTCGCCGGAGGCCATCTCATTGGCGTCCTTCGTC 667

Qy   107 IleIleAlaIleAlaAspVal--------AspGluArgThrArgLysThrGlyGluLeuAlaPheAlaAlaGly 125
         |||:||:|| :::
Db   666 CTCACTGCCGAGGTTACTTGACGAAAAGGGTGAAGGAGGACTCCCTGCCTGCGTGGGGCGCC 607

Qy   126 LeuAlaProAspCysAlaIleThrValHisThrGlnAlaAlaAspThrSerSerProAsp 145
         ||| :::
Db   606 GCCGGCGGTGACTGCGCCATA- 580

RESULT 4

```
US-10-449-902-21879/c
Sequence 21879, Application US20060123505A1
Publication No. US20060123505A1
GENERAL INFORMATION:
  APPLICANT: National Institute of Agrobiological Sciences.
  APPLICANT: Bio-oriented Technology Research Advancement Institution.
  APPLICANT: The Institute of Physical and Chemical Research.
  APPLICANT: Foundation for Advancement of International Science.
  TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
  FILE REFERENCE: MOA-A0205Y1-US
  CURRENT APPLICATION NUMBER: US/10/449,902
  CURRENT FILING DATE: 2003-05-29
  PRIOR APPLICATION NUMBER: JP 2002-203269
  PRIOR FILING DATE: 2002-05-30
  PRIOR APPLICATION NUMBER: JP 2002-383870
  PRIOR FILING DATE: 2002-12-11
  NUMBER OF SEQ ID NOS: 56791
  SOFTWARE: Patentin Ver. 2.1
  SEQ ID NO 21879
  LENGTH: 1470
  TYPE: DNA
  ORGANISM: Oryza sativa
  PUBLICATION INFORMATION:
  DATABASE ACCESSION NUMBER: AK072104
  DATABASE ENTRY DATE: 2001-12-06

US-10-601-319-10 (1-432) x US-10-449-902-21879 (1-1470)

Alignment Scores:
  Pred. No.:
  Score:                0.404
                        106.50
  Percent Similarity:   34.8%
  Percent Identity:     23.2%
  Best Local Similarity: 4.7%
  Query Match:          6

  Length:        1470
  Matches:       70
  Conservative:  35
  Mismatches:    115
  Indels:        82
  Gaps:          13
```

```
Qy   230  rGlyAlaValSerLeuAlaSerMetLeuThrGluIlePheLeuLeuGlnGlnAlaGlnGl  250
               |||
Db   315  CGCCCGACA                                                    306

Qy   250  yMetProGluProGlyTrpGlyArgIleThrAspSer-----------------HisGl  264
              :::|||                           |||||
Db   305  -CTGCGTCCGGCGGCGGTGG-----CGGGACAGGACGGGCCCTGTAGGTAGTGCCACTT  264

Qy   290  ------------AlaThrProLeuLeuAspLeuIle-------------------L  298
                       ::|||    ::|||||||:::
Db   290  AAAATGGCTACGCGCGGGCCCTTCCATTGAATCTCCTGAGGGCATCCAACGTTGCTCTCCACCGA  298

Qy   284  luValAlaArgSerArg-------------------------------------Glu  289
               |||||||||||
Db   195  GCGCGCGCTCGTTACGAGGTGGCGATTGCTCTGCGGACGTTGGCGGAACGCGGAAGATGGT  136

Qy   264  nTrpAsnThrLeuLeu--SerLeuHisAsnLeuGlnPheAspLeuLeuGlnArgThrProG  284
          |||                |||||||||
Db   305  GTGGTCCTCCAAGGTCGCAGTGGGGGCGGCCGCGGCGTCGGTGGTAATAGGGGGCGGATGAA  256

Qy   255  yThrAlaLeuThrProHisPro  305
              |||||||||||
Db    75  AGATTCCAACGAGCGCACATCCA  53

Qy   298  ysThrAlaLeuThrProHisPro  305
          |||||||||||
Db    75  AGATTCCAACGAGCGCACATCCA  53
```

RESULT 5

```
US-10-449-902-7723/c
Sequence 7723, Application US20060123505A1
Publication No. US20060123505A1
GENERAL INFORMATION:
  APPLICANT: National Institute of Agrobiological Sciences.
  APPLICANT: Bio-oriented Technology Research Advancement Institution.
  APPLICANT: The Institute of Physical and Chemical Research.
  APPLICANT: Foundation for Advancement of International Science.
  TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
  FILE REFERENCE: MOA-A0205Y1-US
  CURRENT APPLICATION NUMBER: US/10/449,902
  CURRENT FILING DATE: 2003-05-29
  PRIOR APPLICATION NUMBER: JP 2002-203269
  PRIOR FILING DATE: 2002-05-30
  PRIOR APPLICATION NUMBER: JP 2002-383870
  PRIOR FILING DATE: 2002-12-11
  NUMBER OF SEQ ID NOS: 56791
  SOFTWARE: Patentin Ver. 2.1
  SEQ ID NO 7723
  LENGTH: 2822
  TYPE: DNA
  ORGANISM: Oryza sativa
```

```
Db   754  GTCCGAGATCTCCCCCT------------TGGGGCGACATGGGGGAGGAGCAGCGGTGTCTGGGATGGGGTC  701

Db    68  GGGAGCGGACGAGGGAGGACGAACTTGGAGTAGGAGCTCCACCTTGTCCACGGCCTCGGA  641
          GGTGGGGGATGCGGTTCACCCTCGGTCGGTGGGGGATGGAAGGGGCGGGGGATGAA  592

Qy   102  SerGlyGln--------------  104
          :::
Db    82  HisTyrTrpArgGlnArgLeuValAlaLeuAspGlyLeuLeuProLysCysGlyCysProGln  101
          CACGAACAGCGAGTAGCAGCGAGGTCGGCGACGATGGGACTG--------  593

Qy   121  AlaPheAlaAlaGlyLeuAla------ProAspCysAlaIleThrValHisThrGlnAla  138
          |||||||         :::|||
Db   472  CTCCTCGATGCGGGGTCTCCGGCCCTTCTTGAGGACGTTGTTGGAGGCTGCCGCCGAGAGGCC  413

Qy   139  AspThrSerSerProAsp--------ProLeuPheAsnProLeuLysThrGlyValCys  155
              :::|||
Db   412  GCGCTCCCGCGTGAGGACGACGAGGGCCGACGACCTTCTTGACGCGGGGCGGATGCGCGTGAGGG  353

Qy   156  GlnLeuAsp------AsnAlaAsnValThrArgAspAlaIleLeuGluArgLeuAlaGlyGlySer  173
          :::                       :::|||      :::
Db   352  GAGGTCGATGTCCCTCGACTCTCCTGGTTCACTGTTGGTAGAGCAACCTCCACCAGCGC  293

Qy   105  -----ValAlaIleIleIleAlaAspValAlaSerGlyThrArgLysThrGly-----Glu  120
               ||||||           :::|||
Db   532  GTAGGCGTTGCCCTTCTTGCCCAGCGACGCTTGAGCGCAGTAGCTGACTGCCTTGAG  473

Qy   174  IleAlaAlaPheThrGlyHisThrArgGlnThrAlaPheArg---------------  186
          :::                 :::|||
Db   292  CTCCGAGAAGGGGGCGGGGACCACCACGCGTCCCTGCGCCCGCCGCACCTTGGCCGCCCG  233

Qy   187  ---GluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCysSerLeuLysArgGluLys  205
             :::                   |||
Db   472  CACCAGCTTCATCGCCTC----------CGTCGATCTTCTGCCGTGTTGCGCACCGA  188

Qy   206  GlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeuLysValSerAlaAsp  225
          :::                 |||||
Db   187  GTTCGATACGGCTCCCGCAGCTCACGACGACGACGGCGGCGGGGGCGGGGGGCTGCT  128

Qy   226  CysValSerLeuThrGlyAlaValValSerLeuAlaSerMetLeuThrGluIlePheLeuLeu  245
          :::|||               :::|||
Db    187  GCCGCAGCTAGGAGCCGGAGGCGGAGCGGCGGGGG--- ::: AGTGGCTGG  44

Qy   246  GlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThrAspSerHisGlnTrp  265
          |||      :::
Db   100  ---GGTGCGCCAGGGCGCGAGGAGGAGCCAGGGCGCGGTGGTGGAGAGGTCGGAGCAGGACATGGGCTGG  44
```

```
US-10-601-319-10 (1-432) x US-10-449-902-21879 (1-1470)

Qy    12  LeuIleProLeuThrProGlnSerAlaPheAlaGlnSerGluProGluLeuLysLeuGlu  31
          ::::||::||||                      :::
Db    12  CTTCTCCCGCTCCACGGTCAGCTTCCCTTCCTTGGTGGTGAGCCGGGAA  31

Qy    32  SerValVal-IleValSerArgHisGlyValArgAlaProThrLysAlaThrGlnLeuMe  51
          |||::::                          |||
Db   790  CTCGTCCTCGGTGGGGCGTCCAGCGAC----------ACGGCCGTTGAT  755

Qy    51  tGlnAspValThr-ProAspAlaTrpProThrTrpProValLysLeuGly---------  67
          :::|||               |||
Db    51  ---                                                       67
```

```
;    PUBLICATION INFORMATION:
;    DATABASE ACCESSION NUMBER: AK105071
;    DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-7723

Alignment Scores:
Pred. No.:                          1.02
Score:                              106.50        Length:           2822
Percent Similarity:                 34.8%         Matches:          70
Best Local Similarity:              23.2%         Conservative:     35
Query Match:                        4.7%          Mismatches:       115
DB:                                 6             Indels:           82
                                                  Gaps:             13

US-10-601-319-10 (1-432) x US-10-449-902-7723 (1-2822)

Qy     12 LeuIleProLeuThrProGlnSerAlaPheAlaGlnSerGluProGluLysLeuGlu 31
          ||||||:||| ::|||::: :::|||||:||| |||:||
Db   2306 CTTCTCCCGCTCCACGGTTCAGCTTCCCTTCCTTTCCTTTGGTGTTGGCTGAGCCCGGAA 2256

Qy     32 SerValVal-IleValSerArgHisGlyValArgAlaProThrLysAlaThrGlnLeuMe 51
          ||:|| |||: :::
Db   2255 CTCGTTCCTTCGGTGGCGGTCCACGCAC----------------------ACGCCGTTGAT 2220

Qy     51 tGlnAspValThr-ProAspAlaTrpProThrTrpProValLysLeuGly------- 67
          |||:|:|| ::|:|:: :::|||
Db   2219 GTCGCAGATCTCCCCCT-----TGGGGACCATGGGAGCAGGGTCTGGATGATGGGGTC 2166

Qy     68 ------------GluLeuThrProArgGlyGly-GluLeuIleAlaTyrLeuGly 81
          :::|||:|||:::
Db   2165 GGAGCGGACGAGGGAGGAGACGAACTTGGAGTAGAGGAGCTCCACCTTGTCCACCGCTCGGA 2106

Qy     82 HisTyrTrpArgGlnArgLeuValAlaLeuAspGlyLeuLeuProLysCysGlyCysProGln 101
          ||||||
Db   2105 CACGAACAGCGAGTAGACGAGTCGCAGATGACTG-----------------GGAGTCCTTCAC 2058

Qy    102 SerGlyGln------------------- 104
          |||
Db   2057 GGTGGGGGATGCCGTTCACCTCCACGGTCCGGTGGGGATGAAGGGGGGGGGATGAA 1998

Qy    105 -----------ValAlaIleIleAlaAspValAlaspGluArgThrArgGlyThrGly------- 120
          |||||:|:::|:||||:||| --Glu
Db   1997 GTAGGCGGTTGCCCTTCTTGCCCAGACGACGGGTGTACTCGAGGCCGAGCTTGCTTGAG 1938

Qy    121 AlaPheAlaLeuAlaGlyLeuLeuAla------------ProAspCysAlaIleThrValHisThrGlnAla 138
          |||||:|||:| :::|||
Db   1937 CTCCCTCGATGCCGGGTCTCGGCCTTCCTTGGAGACGTTGTTGAAGCTGCCGCCAGAGGGCC 1878

Qy    139 AspThrSerSerProAsp------------ ProLeuPheAsnProLeuLysThrGlyValCys 155
          :::||| |||
Db   1877 GCGCTCCCCGGGTGAGGACGACGAGGGGCGACCTTCTTGACGGGGGATGGGGATGGGG 1818

Qy    156 GlnLeuAsp----------AsnAlaAsnValThrAspAlaIleLeuGluArgAlaGlyValSer 173
          ||||||:|:::
Db   1817 GAGGTCGATGCGGTCCTCCCTCTGGATCTCCTGGTTCATCGTTGTAGAGAGCACCTCCACCGGC 1758

Qy    174 IleAlaAspPheThrGlyHisTyrGlnThrAlaPheArg----------- 186
          :::
Db   1757 CTCCGAGAAGGGGGCGAGGACACCACCGCTTCCCGCCCCGACCTTGGCCGCCGCCGC 1698

Qy    187 ---GluLeuGlyArgValLeuAsnPheProGlnSerArgAsnLeuCysSerLeuArgGlyLys 205
          :::||| |||
Db   1697 CACCAGCTTCATCGCCTC----------------CGTTGAATCTTCTGCCGTGTTGCGGACGGT 1653

Qy    206 GlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluIleLysValSerAlaAsp 225
          :::
Db   1652 GTCGATACGGTCCCGCAGCTCACGCGACGACCCGACACCTGCAGGCGCCGCCTGCGT 1593

Qy    226 CysValSerLeuThrGlyTyrAlaValSerLeuAlaSerMetLeuThrGlyIlePheLeuLeu 245
          :::||:
Db   1592 GCCGCAGCTAGGAGCGCGAGCGCGGGCG-------------------------- 1566

Qy    246 GlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThrAspSerHisGlnTrp 265
          ||||||:::
Db    246 -----------------------------------------------------
```

```
Db   1565 ---GGTGGGGCCAGGGCCGGAGGAGGAGACCAGCGCGGGGTGGGAGAGGTGCCGAGCACCAGCACATGGGCTGG 1509

RESULT 6
US-11-189-279-64
;    Sequence 64, Application US/11189279
;    Publication No. US20060115829A1
;    GENERAL INFORMATION:
;    APPLICANT: MAO, LI
;    APPLICANT: WANG, JIE
;    APPLICANT: LUO, WANG
;    TITLE OF INVENTION: A METHOD OF TREATING CANCER
;    FILE REFERENCE: UTKC:875US
;    CURRENT APPLICATION NUMBER: US/11/189,279
;    CURRENT FILING DATE: 2005-07-26
;    PRIOR APPLICATION NUMBER: 60/598,554
;    PRIOR FILING DATE: 2004-08-03
;    NUMBER OF SEQ ID NOS: 69
;    SOFTWARE: PatentIn Ver. 2.1
;    SEQ ID NO 64
;    LENGTH: 118899
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-11-189-279-64

Alignment Scores:
Pred. No.:                          211
Score:                              106.50        Length:           118899
Percent Similarity:                 31.1%         Matches:          84
Best Local Similarity:              22.3%         Conservative:     33
Query Match:                        4.7%          Mismatches:       115
DB:                                 7             Indels:           144
                                                  Gaps:             17

US-10-601-319-10 (1-432) x US-11-189-279-64 (1-118899)

Qy     61 ThrTrpProValLysLeuGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
          ||:|| :::||||
Db  23020 ACCTGGCCTTCGCCTTCGGAAAAGTTAAAGGCGCCAGAGAGGAGGAGGTGAG--GAGGGGTTTCTG 23076

Qy     81 GlyHisTyrTrp-ArgGlnArgLeuValAlaLeuAspGlyLeu----------LeuProLy 96
          ||| || ||||:|:
Db  23077 CCATGCAGGTGGGCGTCCGCCCCCTT--GCAACTGGGATGTGGGGTCTTCAGGCGAGCGG 23133

Qy     96 sCysGlyCysProGlnSerGlyGlnValAlaIleIleAlaAspValAlaspGluArgThrAr 116
          |||:::
Db  23134 CTGTCCCTTCCCCCAATCC--------------CACTTTAGGAAAGAGCTCG 23172

Qy    116 glyThrGlyAlaAlaPheAla---AlaGlyLeuLeuAlaProAspCysAlaIleThrValHi 135
          ::|||:::||| |||||:::
Db  23173 TGAGTGTGGGCTTAGACTCGCGCGGGGGGGTC-----TGCCGCCGGTTGGAGGGGGCTT 23223

Qy    135 sThrGlnAlaAspThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValCy 155
          ::||| |||:::
Db  23224 GAGGCAGAGGGGCCGACACCCCCGGGGGTGTCCTGAGCTGAGCGCGGGGGAGCAGC 23283

Qy    155 sGlnLeuAspAsnAlaAsnValThrAspAlaIleLeuGluArgAlaGlyValSerIleAl 175
          :::||||
Db  23284 GGGGGCTGCCCCTCTGCCGTGTGCACACCTGGCCCGGGACGGCCCCGAGGGCGCCCC 23343

Qy    175 aAspPheThrGlyHisTyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnPhePr 195
          ||||||
Db  23344 TCCCTGTCCCGGCCAC---------------- 23359

Qy    195 oGlnSerAsnLeuCysSerLeuArgGlyLysGlnAspGluSerCysSerLeuThrGlnAl 215
          :::
Db  23360 ----------------------GCCGCCTACCCTGGGGT 23376

Qy    215 aLeuProSerGluIleLysValSerAlaAspCysValSerLeuThrGlyTyrAlaValSerLe 235
          :::||:
Db  23377 CTTACCTTCCATTCTTAGCCTAACTCCCGTGTGC-------------------- 23410

Qy    235 uAlaSerMetLeuThrGlyIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGl 255
          ||:::
Db  23411 ---------------------------------------------GG 23412
```

```
Qy    255  yTrpGlyArgIleThrAspSerHis----------------------------------------  263
           |||
Db  23413  CTGGCCCCACTTCCCAGACTCGCACTAAGGCCTCAAAAACCTCTAGCACCAGGTCCTGG       23472

Qy    264  --------------------GlnTrpAsnThrLeuLeuSerLeuHis----------------  272
                               ||||
Db  23473  AAGATAAGACTTCCCCGGCTCGTGGTTGCATGGGAGGGGTGTAATTGAGCTCCACTTTATCGG    23532

Qy    273  ----AsnAlaGlnPheAspLeuLeuGlnArgThrProGluValAlaArgSerAr           289
               |||:::
Db  23533  GGAGAAAGTGGGAGACGCGGAGAAAGGACCTAACAAGTGTGAAGACTGCGCTCCCTTGGGCCC  23592

Qy    289  gAlaThrProLeuLeuAspLeu------------------------------            296
           |||
Db  23593  AGGGCTTCCCGCTGGAGGATCTCACCTGGATGGAGGAGTTCCGAGCCGTCTTGCCCGCG     23652

Qy    297  -IleLysThrAlaLeuThrProHisProPro---------GlnLysGlnIleAlaTyrGlyVa  313
                                                   |||
Db  23653  CCTCAGCCTTCCCTGCTCCTTTTCCTCCCCTGGTCTTAGGAGAGAGGACCAGTTCACGAGATC  23712

Qy    313  l--------ThrLeuProThrSer-----------ValLeuPheIleAlaGlyHi           325
                    |||||||||:::                  :::
Db  23713  CCGTGTTCCACCCTCCCCCGAAGCTCGGGAGTGTTTGAATTAAAGGGGCAG             23772

Qy    325  sAspThrAsnLeuAlaAsnLeu-----------------------GlyGl               334
                                                           |||
Db  23773  AGTATGAAACTTTGAGTTGCCTCCTGAGTTCTCGGATCGGTGGGGTTGCTGCAGACGGGTGG    23832

Qy    334  yAlaLeu--------GluLeuAsnTrpIleLeuProGlyGlyValGlnProAspAsnThrPr   350
           |::|             ||
Db  23833  CAGTTTTTGGGAAACAGCTGTTATCTCCTCTACCAGCCTGCTGGGGGCTTCTTGTTCACC     23892

Qy    350  o--------ProGlyGlyGluLeuValPheGluArgTrpArg     361
                    |||::|||         |||
Db  23893  CACCTTGGAAACGCCAGGTGCCGACTTGAATTTTCAGAGTTGAGA  23938
```

Alignment Scores:

Score:	105.00	
Percent Similarity:	38.1%	
Best Local Similarity:	26.4%	
Query Match:	4.7%	

Length:		1221
Matches:		70
Conservative:		31
Mismatches:		119
Indels:		46
Gaps:		11

```
Qy    106  AlaIleIleAlaAspValAspGluArgThrArgLysThrGlyIleGluAlaPheAlaAlaGly   125
           |||:::                ::::::::::          ::: :::
Db    404  GCTCTCCGGCACCCAGCAACGGAGTCTCAACAGCAGCAGGTGCGGTAGCGTTTCACGGGT     463
```

```
Qy    126  LeuAlaProAspCysAlaIleThrVal-----HisThrGlnAlaAspThrSerSerPro   143
           |||                          :::
Db    464  CTGTCAAG-ACGTGTCGAGGTGCGGCGGTGCCCGGTGCGGGGGACGGGATGCACACACCA    522

Qy    144  AspPro-LeuPheAsnProLeu---LysThrGlyValCysGlnLeuAspAsnAlaAsnva   162
                                     |||
Db    523  CAACCTGCTCTTTTTAAAGGCTGCAGAAAACTTCACTCTGTTAAGAACACATCTGG       582

Qy    162  lThrAspAlaIleLeuLeuArgAlaGlyGlySerIleAlaAspProPheThrGlyHisTyrGl  182
           |||:::                                 :::
Db    583  TATCCCAAATTTAATTTCGACGAAGAGGAATATCCTTCCA-----ACATCACCACTACT     636

Qy    182  nThr--------AlaPheArgGluLeuGlnArgValLeuAsnPheProGlnSerAsnle   199
           |||
Db    637  TACCTCAAGTCGTCGCATTATGATGCTAAAACAAGATCCCTTCGCCCCTATTCCGTCTT    696

Qy    199  uCysLeuValArgGlyGluLysGlnAsp-----------------------------   207
           |||
Db    697  GGCAAAATAGTGGGAGAACGCAGGACACAGTTCCAGGACGATGCGTGGGAGGGGGATC     756

Qy    208  ---GluSerCysSerLeuThrGlnIleAlaLeuProSerGluLeuLysValSerAlaAspCy   226
              |||
Db    757  ATGGGCATCCAGGTCAACTGGGACTGCAACCTGGACAGAGCGCCCTCCCTCTGCT--TG     813

Qy    226  sValSerLeuThrGlyValAlaValSerMetLeuThrGlyIlePheLeuLeuGl           246
           |||
Db    814  CCCAGGTTACTCCTTCCCCGCCTCGATACACGGGAGGTTGAGCACAACGTATCCTCTGGC   873

Qy    246  nGlnAlaGlnIleGlyMetPro-----------Glu-ProGlyTyrTrpGlyArgIleThrAspSerH   263
           |||
Db    874  TACAATTTCAGGTTTGCCAAGTACTACAGAGACGAGCAGCAGCAGCAGCGCTC        933

Qy    263  lsGln-----TrpAsnThrLeuLeuSerLeu-----------------------H    272
           |||
Db    934  ATCAAGGCCTATGGCCATCCGGCTTCGACAATCATTGTGTTGGGGAGGCAGGGAAATTTGAC  993

Qy    272  lsAsnAlaGlnPheAspLeuLeuGlnArgThrProGluValAlaArgSerArgAlaThrP   292
           |||:::
Db    994  ATCAACCCCACTATGATCAACATCGGCTTCTGCGCACTGCTAGGCATGCCCTCCTCG     1053

Qy    292  roLeuLeuAspLeuIleLysThrAlaLeuThrProHisProPro-GlnLysGlnIleAlaTyrG   312
           |||
Db   1054  CA----------CCTTCCCTTCCTCAAGACACCCCCTCAGGTCCCAGCCTTCT        1098

Qy    312  lyValThrLeuProThrSerValLeuPheIleAlaGlyHisAspThrAsnLeuAlaAsnL   332
           |||:::                                 :::
Db   1099  CCCAAGAGTG--------GGAGTGCCCTTCCATTCCGGTAAAAGATTCCAGGCTTCCAGGA  1152

Qy    332  euGlyGlyAla   335
           |||
Db   1153  AGGGGCACGCA  1163
```

```
US-11-145-307A-189

Alignment Scores:
Pred. No.:                1.34          Length:        2711
Score:                  105.00          Matches:         87
Percent Similarity:      34.9%          Conservative:    51
Best Local Similarity:   22.0%          Mismatches:     154
Query Match:              4.7%          Indels:         104
DB:                                     Gaps:            16

US-10-601-319-10 (1-432) x US-11-145-307A-189 (1-2711)

Qy   26 ProGluLeuLysLeuGluSerValValIleValSerArgHisGlyValValArgAlaProThr  45
        |||::::: |||  |:::
Db  239 CCCGGAAGTGGTCCCGCACGCCATCGTAGACCTGGCCACCAGTGATGTAAGGACCCTGCGAC  298

Qy   46 LysAlaAlaThrGlnLeu------------------------------------            46
        :::
Db  299 ACTCTCTGCCAGCTCCCGGCGGTCACCCATGGCCTCAGCAAGANTCCTTGGCAGTACTAC    358

Qy   60 ProThrThrTrpProValLysGluGlyGluLeu------                ThrPro   71
        |||                                                   |||||
Db  359 GAGACTGGACGCAGCATCCGCCTCGAGTGATAGGGGCGCTCCAAGCCCAAGGTGGGCCCC  418

Qy   72 ArgGlyGlyGluLeuIleAlaTyrLeuGlyHisTyrTrpAlaArgGln----------      86
        :::                                    ::: ||
Db  419 AAG----------GTGGTCGGAAGATTGGGGACTACAAACGCCAGAACCCTACCATGTTT    469

Qy   87 ---------------------------ArgLeuValAlaAlaAspGlyLeuLeuProLysGly 100
                                  |||  | ||||::|||||
Db  470 GCCTGGGAGATCCGACACGGCTTCCTGCTGAGCGTCTGTGAACAATGACACTGTGCCC    529

Qy  101 GlnSerGlyIleValAlaAlaIleIleAlaAspValAspArgLysThrArgGlyCysGlu   120
        ::: ||                           |||
Db  530 AGTGTCAGCTCCATTAATAGAATCATC----CGGACC----AAAGTGCAGCAA          574

Qy  121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrVal----------            134
        :::
Db  575 CCATTC------AACCTCCCTATGGACAGCTGCGGTGGCCACCAAGTCCCTGAGTCCCGGA   628

Qy  135 HisThrGlnIleAlaSerThrSerProAlaProLeuPheAsnProLeuLysThrGlyVal   154
        :::|||:::
Db  629 CACACCGCTGATCCCCAGCTCACGCTGTAAACTCCCCGGAGTCACC                673

Qy  155 CysGlnIleLeuAspAsnAlaAsnValThrAspAlaIle---Leu---GluArgGlyGlySe 173
        |||                        |||:::||        :::
Db  674 ---CAGTCGGATTCCCTGGGCTCCACTACTCACTCGGGCTCCTGGGCATCGGCTCAG    730

Qy  187 --------GluLeuGluArgValLeuAsnPheProGlnSerAs                198
        |||                                  :::
Db  731 CCTGGCAGCGGACAAGAGAGGAAAAATGGATGACAGTGATCAGGAATAGCTGCCGACTAAGCATT  790

Qy  198 nLeuCysIleLysTyrGlnAspGluSerCysSerLeuThrGln------              214
        |||                            |||
Db  791 GACTCACAGAGCAGCAGCGACCCCGGAAAGCACCTTCGCACGGATAGCCTTCAGCGAG      850

Qy  215 ------AlaLeuProSerGluLeuLysLysValSerValSerSerLeuThrGlyVal       232
        :::
Db  851 CACCAGCCGGCTCGAGTGCCCATTGAGCGGCAGCACTACCAGAGGCCTAYGCC         910

Qy  232 aValSerLeuThrSerMetLeuThrGluIlePheLeuGluGlnGlnAlaGlyMetPro     252
        :::
Db  911 TCCCCCAGCCACCACAAGGCAGCAGGCCTCTACCGGCCTCTACCGGTTGCTCAACAGCACC   970

Qy  252 oGluProGlyThrGlyArgIleThrAspSerHisGlnTrpAsnThrIleLeuSerLeuHi   272
        :::|||||::
Db  971 CTGGACGACGGGA----------AGGCACCCTGACCCCTTC                    1003

Qy  272 sAsnAlaGlnPheAspLeuLeuGlnArgThrProGluValAlaAlaArgSerArgAlaThrPr  292
        |||:::||
Db 1004 AACACGCCACTGGGGGCGGCCAACCTCTCG-ACTCACCAGACCTACCCCTGGTGGCCAGATCC  1062
```

```
Db 1063 TCACTCACCCTTGGCCATAAAGCAGGAAAACCCCGGAGGTGTCCCAGTTCTTCCACCCC  1122
        :::||::
Qy  305 oProGlnLysGlyIleLysValThrLeuProThrSerValLeuAsnPhe-------     321

Db 1123 TTGCTCTTTATCTAGCTCCCGCCCTTTTGGATCTG----------------          1156
                    ----LeuLeuAsnPheLeuIleLeuLysThrAlaLeuHisProHisPr  305

Db 1207 CTCCGGTGTACGGGGCAGTTCACGGGGCCAGGGGCCCTCCTTCAGGGCGGGGAGGATGGTGGGGCC  1266
        |||                                    |||::: |||
Qy  322 ----IleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTyr  340

Db 1267 CACGCTGCCCGGGATACCCCACACCCCACACATCCCACACCAGCGGCGACAG          1309
        :::                                    :::
Qy  340 pThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlu                   354

RESULT 9
US-10-501-834-1/c
; Sequence 1, Application US/10501834
; Publication No. US20060088828A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Peter C., Ward, Christopher J., Rossetti, Sandro, and Torres,
; APPLICANT: Vicente E.
; TITLE OF INVENTION: Polycystic Kidney Disease Nucleic Acids
; TITLE OF INVENTION: and Proteins
; FILE REFERENCE: 07039/386US1
; CURRENT APPLICATION NUMBER: US/10/501,834
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/US03/02038
; PRIOR FILING DATE: 2003-01-23
; PRIOR APPLICATION NUMBER: 60/351,110
; PRIOR FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 221
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 12225
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-501-834-1

Alignment Scores:
Pred. No.:                12.8          Length:        12225
Score:                  104.50          Matches:         109
Percent Similarity:      34.9%          Conservative:     54
Best Local Similarity:   23.3%          Mismatches:      167
Query Match:              4.6%          Indels:          139
DB:                                     Gaps:             24

US-10-601-319-10 (1-432) x US-10-501-834-1 (1-12225)

Qy    1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla   20
        |||||
Db 4439 ATGAAAAGAGTGCAAATTCCCCTGACACTCGCTGGTTAGCCCATTGACACAGGACGTGTGACG  4380

Qy   21 PheAlaGlnSerGluProLeuLeuLysGluSerValValIleValSerArgHisGly    40
        |||  |||
Db 4379 TTCAGGGAGGAAGGAAGGCTCCA----------                        4356

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro  60
        :::|||||::                              |||
Db 4355 AAGGGGTCACCCTCACCGGCCTCCCTTGCTCACCAGACC-                  4326

Qy   61 --ThrTrpProValLysGluGlyGluLeuThrProArgGlyGlyGlu----          75
        :::                                  |||
Db 4325 AGAATGGGTGTGGTCTCCCAAAACTCACACAGTAAAAGGACACCCGAGAGGTCACC  4266

Qy   76 ------LeuIleAlaTyrLeuGlyHisTyrTrpArgGlnArgLeuValAlaAspGly  92
        :::|||::
Db 4265 CGAACTGACCTCCTTCTAGAGTTA----------AGAAGCAACCCCTCCACAGTAAGT  4218
```

RESULT 10
US-10-511-937-358

; Sequence 358, Application US/10511937
; Publication No. US20060088836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 50661200000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 358
; LENGTH: 3646
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-511-937-358

US-10-601-319-10 (1-432) x US-10-511-937-358 (1-3646)

Alignment Scores:
Pred. No.:
Score: 4.37
Percent Similarity: 101.50
Best Local Similarity: 32.9%
Query Match: 22.3%
DB: 4.5%

Length: 3646
Matches: 124
Conservative: 59
Mismatches: 195
Indels: 181
Gaps: 29

Db   3304 ATGTAAAATGCTCTGGGA 3288

```
RESULT 11
US-11-293-697-315

; Sequence 315, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: H1-A0106
;  CURRENT APPLICATION NUMBER: US/11/293,697
;  CURRENT FILING DATE: 2005-12-05
;  PRIOR APPLICATION NUMBER: US/11/108,260
;  PRIOR FILING DATE: 2002-03-28
;  NUMBER OF SEQ ID NOS: 5458
;  SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 315
;  LENGTH: 2408
;  TYPE: DNA
;  ORGANISM: Homo sapiens
  DB: US-11-293-697-315

Alignment Scores:
Pred. No.:                4.17
Score:                   99.00              Length:        2408
Percent Similarity:      30.5%              Matches:         65
Best Local Similarity:   23.0%              Conservative:    21
Query Match:              4.4%              Mismatches:     106
DB:                                         Indels:          90
                                            Gaps:            13
```

US-10-601-319-10 (1-432) x US-11-293-697-315 (1-2408)

```
Qy   5  LeuLeuProPheLeuSerLeuLeuIle------------------ProLeuThrProGlnSerAla  20
        |||:|||||||||                                ||  |||:|||
Db 255  CTGCTTCCCAGAGTCTCCCGACCCAAAATTCCTCTCCTCAAGACCGACGGCGCGACGCTGC      314

Qy  21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValAlaSerArgHisGly  40
        :::                       |||||||||
Db 315  AAAAGCTCGTCCCCGGCCTCACCTG-                                         347

Qy  41  ValArgAlaProThrLysAlaThrArgGlnLeuMetGlnAspValThrProAspAlaTrpPro  60
        |||:|||    ||||||              |||||||||||||||
Db 348  CGGCGCGTCCCCTTGAAGGCGGCGGGC-----CAGGATCGGCGGGTCTCGCCGGCCTGGGGC    401

Qy  61  ThrTrpProValLysLeuGlyLeuThrProAsnThr-----GlyGluLeuIleAla          78
        :::       |||                            ||:|||
Db 402  GTTTGGCTTTTGTCCCGCGCCCGAGGTCCGGCGCCCAGGAGAGCTCCCGGGC              461

Qy  79  TyrLeuGly----------------------HisTyr                             83
        |||||||||                      :::
Db 462  GGCCTCGGGGACAAAAGGCCCGGCGACGTGCAGAGTTCACCCAGAGGTCCG                521

Qy  84  TrpArgGlnArgLeuValAlaAspGlyTyrLeuCysCysProGlnSerGly             103
        |||:|||     |||:||||
Db 522  ACCAGAAAACGAGAGAACGCCCGGAGGGAGGAGGCGGCTGC-----                   569

Qy 104  GlnValAlaIleIleAlaAspValAspGluArgArgLysThrArgGlyValAlaPheAla    123
        :::|||||||              :::|||||||        |||:|||
Db 570  TGGCTGGGAGAACCGACGCGGCAACCAGACGGGCAAAGCCGGAAAGG                  617

Qy 124  AlaGlyLeuAlaProAspAspCysAlaIleIleThrValHisThrGlnAlaAspThrSerPro 143
        |||||||     :::|||||||         |||:||||
Db 618  GCAGGCGCGAGGACC----------------------ACCTGACGCGCGCTCTCCCAGCCGGCTCTCCT 665
```

```
Qy  93  LeuLeuProLysCysGlyCysProGlnSerGlyGlnIleValAlaIleIleAlaIleAlaAspValAlaAsp 112
        |||:::                                              |||:::
Db 832  ---ATCTCCCTCGGTGGCTCCTCGGGTTCCTCCCGTGACTGTGACCTG---                      885

Qy 113  GluArgThrArgGlyLysThrGlyValAlaPheAlaAlaGlyLeuAlaProAspCysAlaIle   132
        ||||||      :::||||||
Db 886  --CCTGGAGATGCAGTGGCACTTGGACTG-------                                     912

Qy 133  ThrValHisThrGlnAlaAspThrSerSerProLeuPheAsnProLeuLysThr            152
        |||
Db 913  ----CAATGCTTTACC----                                                     924

Qy 153  GlyValCysGlnLeuAspAlaAsnAlaAsnValThrAspAlaLeuGluArgAlaGlyGly     172
        :::|||
Db 925  ----------TTGGACCTGAAGAAATGTTACCTGTCAATGGCAGCAACCACCATGGCT        972

Qy 173  SerIleAlaAspPheThrGlyHisThrTyrGlnThrAlaPheAlaArgGluLeuGluArgValLeu 192
        |||:::     |||
Db 973  AGCTCCCAAGGCTTCTTACCACAGCAGGCACGGTTGCCCCAGAGACAGG------          1026

Qy 193  AsnPheProGlnSerAsnLeuCysLeuLysArgGluLys                          205
        |||     :::|||||||
Db 1027 --TACCCCATCTGGGAGGAACTGGACAGGAGAAACAAATCCAGGACTACAGACC          1083

Qy 206  ---GlnAspGluSerCysSerLeuThrGlnLeuAlaLeuProSerGluLeuLysValSerAla 224
        :::|||||||
Db 1084 CCACAGTTCTTCTGCTGCCCACTTCAAGTCACAGAAATGACAGCATTATTCACATCCTTGTG 1143

Qy 225  AspCysValValSerLeuThrGlyValAlaValSerMetLeuThrGluIlePheLeu       244
        :::||||||
Db 1144 GAGGTGACCACCAGCCCCGGGTACTGTTCAC------                           1197

Qy 245  LeuGlnGlnIleAlaGlyHisGlyMetProGluProGly------TrpGlyArgIleThrAspSer 262
        |||||||||||:||||||             |||||||
Db 1198 ATCCACCAGGCTGTGCGCCCTCCCACCCCAAACTTGCACTGGAGGGAGGAGGATCTTCCAGTGGG 1257

Qy 263  His--------GlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeu     279
        |||     :::||||||
Db 1258 CATCTGGAATTGGAGTGGCAGCAC----CCATCGTCCTGGGCAGCCCAAGA-GACCTGTTA  1313

Qy 280  GlnArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuLeuAspLeuIleLysThr 299
        ||||||      :::|||
Db 1314 TCA----ACTCCG--ATACACGGAGAAGGGCCATCAGGAGCTGGAAGTGCTGGAGCCGGCC    1367

Qy 300  AlaLeuThrProHisProGlnLysGlyGlnIleAlaTyrGlyValThrHisLeuProThrSerVal 319
        |||:::
Db 1368 TCTCGGGGGCCCGGAGGGGGAGAAGGGGGGACCCTGGACTGCGCCGGCGGGATCTGCTACCG--- 1418

Qy 320  LeuPheIleAlaGly------------HisAspThrAsnLeuAlaAsnLeuGly          333
        |||
Db 1419 ---TTTACAGCTGCGCGCCGCCAGGCTCAACGGGCCCCACCT-ACCAAGGTCCCCTGGAGCTCGGT 1474

Qy 334  GlyAlaLeuGlnLeuAsnTrpProThrLeuProGlyGlnProAspAsnThrPro------    350
        |||:::
Db 1475 GGTCGGACCCAACTAGGGTGGGAGACCGGCCTGGGATCCTTGGTGACCG               1534

Qy 351  ------------------ProGlyGlyLeuValPheGluArgTrp---                360
        |||
Db 1535 CTCTGCAATCTAGTGCTGGGCCGTCCTCAGCGCCGTCCTG-GGCCTGCTGCGCTGGAGGTGGCAG 1593

Qy 361  ------------------ArgArgLeu                                     363
        |||||||:::
Db 1594 TTTCCTCCACACTACAGGACTGAGGCATGCCCTGTGGCCCTGCACTTCCAGACCTGCAC     1653

Qy 363  ------------------                                             363
Db 1654 CGGGGTCCTAGGCCAGTACCTTAGGGACTCAGCCCTGCCAAGGCCACAGTC            1713

Qy 364  SerAspAsnSerGlnTrpIleGlnThrLeuGlnIleMetArg                     383
        |||     :::      :::
Db 1714 TCAGATACCTGTGAAGAAGTGGAACCCAGCTTCCTT--GAAATCTCCCAAGTCCTCA       1770

Qy 384  AspLysThrProLeuSerLeuAsnThrProProGlyGluVal----                 397
        |||||||:::
Db 1771 GAGAGGACTCCTTTGCCCCTGGTTCCTCCAGGCGCCAGAGCCCAGATGGACTACCGAAGATTGCAG 1830

Qy 398  ------------------LysLeuThrLeuAlaGlyLeuCysGluGlyLeuArgAsnAlaLeuGlyLysMet 412
        :::
Db 1831 CCTTCTTGCCTGGGCGGACCACGCCCCTGTCTGTGTGCCCACCCATGGCTCAGGGTCC     1890

Qy 413  CysSerLeuAlaGlyPheThrGlnIleIleValAlaAsnGluAlaArgIlePro         428
        |||     :::     |||
Db 1891 TGCTGTACC--------ACCACAATTGCCAACCAATTCCTACCTACCA                1929
```

RESULT 12
US-10-857-260-27
; Sequence 27, Application US/10857260
; Publication No. US20060110742A1
GENERAL INFORMATION:
APPLICANT: Lyons, Leslie A.
APPLICANT: Grahn, Robert
APPLICANT: Erdman, Carolyn
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: Carrier Tests for Polycystic Kidney Disease in the Cat
FILE REFERENCE: 023070-14680US
CURRENT APPLICATION NUMBER: US/10/857,260
CURRENT FILING DATE: 2004-05-28
NUMBER OF SEQ ID NOS: 47
SOFTWARE: Patentin Ver. 2.1
OTHER INFORMATION: GenBank Accession No. AC145332.28 genomic DNA for
OTHER INFORMATION: PKD1 39536 bp contig
SEQ ID NO 27
LENGTH: 39536
TYPE: DNA
ORGANISM: Felis catus
FEATURE:

Alignment Scores:

Pred. No.:		
Score:	251	
Percent Similarity:	98.50	Matches: 39536
Percent Identity:	34.0%	Conservative: 93
Best Local Similarity:	24.3%	Mismatches: 37
Query Match:	4.4%	Indels: 149
DB:	6	Gaps: 103
		Length: 20

US-10-601-319-10 (1-432) x US-10-857-260-27 (1-39536)

RESULT 13
US-11-317-329-2
; Sequence 2, Application US/11317329
; Publication No. US20060105413A1

GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
; FILE REFERENCE: UM-06967
; CURRENT APPLICATION NUMBER: US/11/317,329
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/128,174
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2696
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-317-329-2

Alignment Scores:
Pred. No.:

	Length:	2696
Score:	8.44	
	Matches:	104
Percent Similarity:	96.50	
	Conservative:	53
Percent Identity:	33.1%	
	Mismatches:	167
Best Local Similarity:	21.9%	
	Indels:	152
Query Match:	4.3%	
	Gaps:	22
DB:	7	

US-10-601-319-10 (1-432) x US-11-317-329-2 (1-2696)

Qy   8 PheLeuSerLeuLeuIleProLeuThrProGlnSerAlaPheAlaGlnSerGlu----------- 25
          |||||||
Db 1207 TTTCTAGACAATTACCATCGAGACAAGACATTACTGCTGTCACTGTCAGAGTCGTGTGGCA 1266

Qy 26 ---ProGluLeuLysLeuGluSerValValIleValSerArgHisGlyValArgAlaPro 44
         ||||||
Db 1267 GTCCCAGAGCTGAAGAGCAAGGCCCTGGCCAGGAAGGTGTCCGAAGCTGTGCGGCCAGCC 1326

Qy 45 ThrLysAla----------ThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
      :::
Db 1327 AGGGAGGGTTAATGAGACAATCAGCCAGGAACTGATGCAGGACAGTGACTCAGGAAACTACTG 1386

Qy 61 ThrTrpProValLysProValLeuGlyGlu-------LeuThrProArgGlyGlyGluLeuIle 77
     ::: |||
Db 1387 AAGCGGGGCCCTTCAGCCTCTTCCGACACCGTTAAGAATTTGGTCCGGAGGATGAAGAACTGTGT 1446

Qy 78 AlaTyrLeuGlyHisTyrTrp---ArgGlnArgLeuValAlaAspGlyLeuLeuProLys 96
     |||||||| :::
Db 1447 AACTATGAGACAACAAGGTCACCCCCCCCAAATTCCTGGTGGCCCAGGGCAGTGTGGAG--- 1503

Qy 97 CysGlyCysProGlnSerGlyGlnValAlaIleIle----------AlaAspValAsp 112
     :::::::|||
Db 1504 CAGGTGAGGTTGCTGCTGCTGGGCCCACGAGTAGACGTGGAC 1542

Qy 113 GluArgGlyThrArgLysThrGlyIleAlaGlyLeuAlaPro---------- 128
     ||||||
Db 1543 TGCCAGACG----------GCCTCTGATACAGACGCCCCTCCTTCCTTCCTGATCGCC 1581

Qy 129 -----------AspCysAlaIleAlaThrValHisThrGlnAlaAspThrSer 141
     ||||||
Db 1582 GCCCAGGACGACCAGCCAACCGACCTTCTGTGCGCCTCTTTGGCACATGGTGCTGCTGATGCCAAC 1641

Qy 142 SerProAsp-----------ProLeuPheAsnProLeuLysThrGlyValCysGln 156
     ::: |||||
Db 1642 CGAGTGGATGAGGATGGGTGGGCGCCCCACTGCACTTTGCAGCCCAGAATGGGATGACCCGC 1701

Qy 157 -----------LeuAspAsnAlaAsnValThrAspAlaIleLeuGluArgAlaGly 171
     |||
Db 1702 ACTGCGCGCGCCCTGCTTCCTGGACACACGGCGGGCGCCTGTGTCAGCAGG 1758

Qy 172 GlySerIleAlaAspPheThrGlyHisTyrGlnThrAlaPheArgGluLeuGluArgVal 191
     ::: ||| |||
Db 1759 TGGACCCCTCTTCACCTGGCCTGCA----------CAGAATAACTTTGAGAATTGGCACGGCTT 1812

Qy 192 LeuAsnPheProGlnSerAsnLeuCysLeuLysArgGlyLysGlnAspGluSerCysSer 211
     ||| :::::::
Db

Db 1813 CTGCTTCCGGTCAGGTGACCCCAACCTGCGGTGAG----------- 1848

Qy 212 LeuThrGlnAlaLeuProSerGluLeuValGlyValSerArgAlaAspCysValSerLeuThrGly 231
     :::||| ||| :::
Db 1849 -----------GCTGAGGGCAGAGACCCCCTCCATGGCCGCTGGCCGCTACTTTGGC 1890

Qy 232 AlaValAlaSerLeuAlaSerMetLeuThrGluIlePheLeuLeuGlnGlnAlaGlyGlyMet 251
     |||
Db 1891 CATGTTAGCCTGGCTGAAGCTGCTGACC----------- 1917

Qy 252 ProGluProGlyTyrTrpGlyArgIleThrArgProSerHisGlnThrpAsnThrLeuLeuSerLeu 271
     || |||
Db 1918 -----------AGCCAGGGGGCTGAGTTGGATGCTCAGCAGCAGAAAACCTGAGAACCACCACTG 1968

Qy 272 HisAsnAlaGlnPheAspLeuLeuGlnAsnArgThrProGluValAlaLeuArgSerArgAlaThr 291
     ||| :::
Db 1969 CACCTGGCA-----------GTAGAGCGGGGCGCAAAGTGAGG 1998

Qy 292 ProLeuLeuAspLeuIleLysThrArgProHisProGluLysIleGlnAlaThrTyr 311
     ::: ||||
Db 1999 GCCATCCAACACCTTGCTGAAGGAGTGGAGCGGTCCCTGGAAGAGTGGGTGCT 2058

Qy 312 Gly---------- 312
     |||
Db 2059 GGCCCACTGCACACTGCAGCTGCCACGGGGCGCAAATACCTGATACTGCAAGATGCTGCTCAGG 2118

Qy 313 -----------ValThrLeuProThr-----------SerValLeuPheIleAlaGly 324
     ||||||||||| :::|||
Db 2119 TACGGAGCCAGCCTTGAGCTGCCCACCCACCCAGGGCTGACACCCCTCGATCTACAGCC 2178

Qy 325 HisAspThrAsnLeu-----------AlaAsnLeuGlyGly 334
     :::
Db 2179 TACAAGGGCCCACTGGAGATCATCCATCTGCTGCAGAGAGCCACGCAAACATGGGTGCT 2238

Qy 335 AlaLeuGluLeuAsnTrpThrLeuProGlnProAspAsnThrProProProGlyGlyGlu 354
     :::|||||||||| |||
Db 2239 CTTGGAGCTGTGAACTGGACCCTCC---CCTAGCTGCCAGCGCCACGGGGAGGA--- 2291

Qy 355 LeuValPheGluArgTrpArgLeuSerAspAsnSerGlnTrpIleGluIleValSerLeu 374
     :::
Db 2292 -----------GGGGTGGTGTCAGCACTGCTGCA 2315

Qy 375 ValPhe-----------GlnThrLeuGlnMetArgAspLysThrProLeuSerLeuAsn 391
     ::: :::|||
Db 2316 GTGTGGGCGTGACCCCAATGCTCAGGCTGAGACCAGACACCCCT----------- 2363

Qy 392 ThrProProGlyGluValLysLeuThrLeuAlaGlyCysGluGlu 406
     ::: ||||||||
Db 2364 ---CCACCTGGCGGGTCCAGGAGGAGCACCTTCCTGGAGTGTCATCAA 2405

RESULT 14
US-11-317-329-4
; Sequence 4, Application US/11317329
; Publication No. US20060105413A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
; FILE REFERENCE: UM-06967
; CURRENT APPLICATION NUMBER: US/11/317,329
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/128,174
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2696
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-317-329-4

Alignment Scores:
Pred. No.:   8.44   Length: 2696

Left column

Score: 96.50
Percent Similarity: 33.1%
Best Local Similarity: 21.9%
Query Match: 4.3%

Matches: 104
Conservative: 53
Mismatches: 167
Indels: 152
Gaps: 22

DB:

US-10-601-319-10 (1-432) x US-11-317-329-4 (1-2696)

```
Qy      8 PheLeuSerLeuLeuIleLeuIleProLeuThrProGlnSerAlaPheAlaGlnSerGlu------  25
            |||||                                              :::
Db   1207 TTTCTAGACATTACCATCGAGACAGACATACTGCTGCAGAGTCGTGTGGCA                 1266

Qy     26 ---ProGluLeuLysLeuGluSerValValIleValSerArgHisGlyValArgAlaPro         44
            ||||                        :::                |||||
Db   1267 GTCCCAGAGACGAAGGCCCTGGCCAGGAAGGTGTCCTGCAAGCTGTCGCCTGGCCAGCCC        1326

Qy     45 ThrLysAla------------ThrGlnIleMetGlnAspValThrProAspAlaTrpPro         60
            :::                   :::             |||
Db   1327 AGGGAGGTTAATGAGGACATCAGCCAGGAAACTGATGGACAGTGACTCAGGAAACTACTG        1386

Qy     61 ThrArgProValLysLeuGlyGlu-----------LeuThrProArgGlyGlyLeuIle          77
            |||            :::                  |||
Db   1387 AAGCGGCCCTTCAGCTTCCCGACCGTAAGAATTTGGTCCCAGACGAAGAACTGTGT           1446

Qy     78 AlaTyrLeuGlyHisTyrTrp----ArgGlnArgLeuValAlaAspGlyLeuLeuProLys      96
            |||                     |||||        :::
Db   1447 ATCTATGAGACAACAAGGTCACCCCCCTCCAATTCCTGGTTGGCCCCAGCCAGTGGGAG---     1503

Qy     97 CysGlyCysProGlnSerGlyGlnValAlaIleIle---------------AlaAspValAsp    112
            :::                     :::                       |||||
Db   1504 CAGGTGAGGGTTGCTGCTGGGCCCCACGGTAGACGTGGAC                          1542

Qy    113 GluArgThrArgLysThrGlyLeuAlaGlyLeuAlaPro-------                     128
            |||                        |||
Db   1543 TGCCAGACG                                                         1581
```

Right column

RESULT 15
US-11-317-329-5
; Sequence 5, Application US/11317329
; Publication No. US20060105413A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
; FILE REFERENCE: UM-06967
; CURRENT APPLICATION NUMBER: US/11/317,329
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/128,174
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 44
; SEQ ID NO 5
; SOFTWARE: Patentin version 3.1
; LENGTH: 2696
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-317-329-5

Alignment Scores:
Pred. No.:   8.44
Score: 96.50
Percent Similarity: 33.1%
Best Local Similarity: 21.9%
Query Match: 4.3%

Length: 2696
Matches: 104
Conservative: 53
Mismatches: 167
Indels: 152
Gaps: 22

DB:

US-10-601-319-10 (1-432) x US-11-317-329-5 (1-2696)

```
Qy      8 PheLeuSerLeuLeuIleLeuIleProLeuThrProGlnSerAlaPheAlaGlnSerGlu------  25
            |||||                                              :::
Db   1207 TTTCTAGACATTACCATCGAGACAGACATACTGCTGCAGAGTCGTGTGGCA                 1266

Qy     26 ---ProGluLeuLysLeuGluSerValValIleValSerArgHisGlyValArgAlaPro         44
            ||||                        :::                |||||
Db   1267 GTCCCAGAGACGAAGGCCCTGGCCAGGAAGGTGTCCTGCAAGCTGTCGCCTGGCCAGCCC        1326

Qy     45 ThrLysAla------------ThrGlnIleMetGlnAspValThrProAspAlaTrpPro         60
            :::                   :::
Db   1327 AGGGAGGTTAATGAGGACATCAGCCAGGAAACTGATGGACAGTGACTCAGGAAACTACTG        1386
```

```
Qy    61  ThrTrpProValLysLeuGlyGlu----------LeuThrProArgGlyGlyGluIle  77
                                           ::: :::
Db  1387  AAGCGGGGCCCTTCAGCTCTCCGACCGTAAGAATTTGGTCCGGACGGAGGANGAACTGTGT  1446

Qy    78  AlaTyrLeuGlyHisTyrTrp---ArgGlnArgLeuValAlaAlaGlyLeuLeuProLys  96
                                 |||||||||||||||||      :::
Db  1447  ATCTATGAGAACAANAGGTACACCCCCCCCCCAATTCCTGGTGGCCCAGGGAGCAGTGTGGA---  1503

Qy    97  CysGlyCysProGlnSerGlyGlnValAlaIleIle------------AlaAspValAsp  112
                                                         :::
Db  1504  ---------------CAGGTGAGGTTGCGTGCTGGGCCCACGAGGTAGACGTGGAC  1542

Qy   113  GluArgThrArgLysThrGlyGluAla------LeuAspAsnAlaAsnValThr----Pro------  128
              |||:||          |||:||:::                :::
Db  1543  TGCCAGACG---------------GCCTCTGGATACAGCCCCTCCTTGAATCGCC  1581

Qy   129  ----------AspCysAlaIleThrValHisThrGlnAlaAspThrSer  141
                      |||||||||        :::
Db  1582  GCCCAGGACCAGCAAACCCGACCTCGTGCCCTGCTTTGGCACATGGGTGCTGATGCCAAC  1641

Qy   142  SerProAsp-----------------ProLeuPheAsnProLeuLysThrGlyValCysGln  156
                                       |||:::
Db  1642  CGAGTGGATGAGGAATGGCTGGGGCCCCACTGCACTTTGCAGCCCCAGAANTGGGANGACCGC  1701

Qy   157  --------------LeuAspAsnAlaAsnValThrAspAlaIleIleLeuGluArgGlaGly  171
                          |||||||:::        |||||||||
Db  1702  ACTGCGCGGGCCCTGCTCCTGGACCACGGGCCCTGTGTGGATGCC---CAGGAACGTGAAGGG  1758

Qy   172  GlySerIleAlaAspPheThrGlyHisTyrGlnThrArgLeuAlaPheArgGluLeuArgVal  191
                                       |||||||||                 :::
Db  1759  TGGACCCCTTCTTCACCTGGCTGCA-----CAGAATAACTTTGAGAANTGGCACGGGCTT  1812

Qy   192  LeuAsnPheProGlnSerAsnLeuCysLeuLysArgGlyLysGlnAspGluSerCysSer  211
                                       |||
Db  1813  CTGGTCTCCCGTCAGGCTGACCCCAACCTGCGTGAG-------------------  1848

Qy   212  LeuThrGlnAlaLeuProSerGluLeuLysValAlaSerAlaAspCysValSerLeuThrGly  231
              :::        |||||||        :::
Db  1849  ---------------GCTGAGGGCAAGACCCCCCTCCATGTGGCGCCCTACTTNGGC  1890

Qy   232  AlaValSerLeuAlaSerMetLeuThrGluIlePheLeuLeuGlnGlnAlaGlnGlyMet  251
                |||||||||||                     :::
Db  1891  CATGTTAGCCTGGCGTCAAGCTGCTGACC------  1917

Qy   252  ProGluProGlyTrpGlyArgIleThrAspSerHisGlnThrAsnThrLeuLeuSerLeu  271
                                 |||        :::
Db  1918  -----AGCCAGGGGGCTGAGTTGGATTGCTCAGCAGAGAAACCTGAGAACAACCACTG  1968

Qy   272  HisAsnAlaGlnPheAspLeuLeuGlnArgThrProGluValAlaArgSerArgAlaThr  291
                |||||||||                 :::
Db  1969  CACCTGGCA---------------GTAGAGCGGGGCAAAGTGAGG  1998

Qy   292  ProLeuLeuAspLeuLysLysThrAlaLeuThrProHisProProGlnLysGlnAlaTyr  311
                  |||||||||:::||:::                |||
Db  1999  GCCATCCAACACCTGCTGCAAGACGTGGAGGGGTCCCTGATGCCCTGACGAGCGGGGCTAT  2058

Qy   312  Gly-----------------------------------------------------------  312
              |||
Db  2059  GGCCCACTGCACACTGCGAGCTGCCAGGGGCAAATACCTGACATCTGCAAGGATGCTGCTCAGG  312

Qy   313  --------------ValThrLeuProThr-----------------SerValLeuPheIleAlaGly  324
                          :::        |||                   :::
Db  2119  TACGGAGCCAGCCTTGAGCTGCCCACCCACCCTGGACACCCCGCATCTAGCAGCC  2178

Qy   325  HisAspThrAsnLeu-----------------------AlaAsnLeuGlyGly  334
                |||                               |||||:::|||
Db  2179  TACAAGGGGCCACCTGGAGATCATCCACTCTGCTGGGCAGAGAGCCACGCAAACATGGGGCT  2238

Qy   335  AlaLeuGluLeuAsnTrpThrLeuProGlyProAspAsnThrProProGlyGlyGlu  354
                |||||||||||||||||                         |||||||||:::
Db  2239  CTTGGGAGCGTGTGAAACTGGACTCC---CCTAGCTGCAGCGCCACGCGGCGGGGAGGA---  2291

Qy   355  LeuValPheGluArgTrpArgLeuSerArgAspAsnSerGlnTrpIleGluGlnValSerLeu  374
                :::
Db  2292  ----------------GGCGGTGGGTGTGTCAGCACTGCTGCA  2315

Qy   375  ValPhe----------GlnThrLeuGlnMetArgGlyAspLysThrProLeuSerLeuAsn  391
              |||:::            :::        |||||||:::
Db  2316  GTGTGGGGGGGCTGACCCCAATCGTCGAGACGAGCTGGACACCCCT----------  2363

Qy   392  ThrProProGlyGluValLysValLeuThrLeuAlaGlyCysGluGlu  406
              |||||||||||||        :::
Db  2364  ---CCACCTGGGCGGTTCCAGAGAGGACACCTTCCTGAGGTGTCATCAA  2405
```

THIS PAGE LEFT BLANK

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

Run on: June 14, 2006, 11:48:02 ; Search time 4895.67 Seconds
(without alignments)
7401.587 Million cell updates/sec

Title: US-10-601-319-10
Perfect score: 2258
Sequence: 1 MKAILIPFLSLIILPLTPQSA.........CSLAGFTQIVNEARIPACSL 432

Scoring table: BLOSUM62

	Xgapop	10.0	,	Xgapext	0.5
	Ygapop	10.0	,	Ygapext	0.5
	Fgapop	6.0	,	Fgapext	7.0
	Delop	6.0	,	Delext	7.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/abs0/ABSSWEB_spool/US10601319/runat_12062006_180052_27516/app_query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cd1 -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs803p
-USER=US10601319 @CGN 1 1 12067 @runat_12062006_180052_27516 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*

				Query		
			Score	Match Length	DB	ID
c	5:	gb_est1:*	847	37.5	595	5	CP326092	JMT1-05-
c	6:	gb_est3:*	839	37.2	895	8	CN762997	ID0AAA5DC
c	7:	gb_est4:*	778	34.5	868	8	CN754382	ID0AAA12D
c	8:	gb_est5:*	768	34.0	864	8	CN757565	ID0AAA1DF
c	9:	gb_htc:*	623	27.6	383	7	AW036132	EST274508
c	10:	gb_est2:*	623	27.6	383	7	AW036134	EST274510
c	11:	gb_est7:*	621	27.5	393	7	BE520240	M11B12STM
c	12:	gb_est8:*	583.5	25.8	531	3	BU713769	SJAABUG01
c	13:	gb_est9:*	522	23.1	706	3	CN759004	ID0AAA24B
c	14:	gb_est0:*	483	21.4	354	1	AA545747	HBMSP1B4-
c	15:	gb_est6:*	443	19.6	707	2	BJ619443	BJ619443
c	16:	gb_htc2:*	389.5	17.2	1175	9	CX943192	LamjGestG
c	17:	gb_est2:*	368	16.3	320	7	AW036134	EST407481
c	18:	gb_est3:*	365.5	16.2	821	3	BE436403	ID0AA1d_E
c	19:	gb_est1:*	360.5	16.0	842	13	CL665745	PRI014c_E
c	20:	gb_est4:*	335	14.8	208	2	BM409016	EST583343
c	21:	gb_est5:*	335	14.8	208	2	BM412806	EST587133
c	22:	gb_est6:*	328	14.5	546	2	BJ040796	BJ040796
c	23:	gb_est7:*	298	13.2	657	4	CA093061	SCCCL200
c	24:	gb_est8:*	288	12.8	657	2	BG457132	NF100C0SP
c	25:	gb_est9:*	284	12.6	1171	8	CV672454	RET7SJ_01
c	26:	gb_est0:*	267	11.8	179	7	BE461872	EST413Z91
c	27:	gb_est1:*	258	11.4	508	2	BJ030203	BJ030203
c	28:	gb_est3:*	239.5	10.6	480	1	AL586973	AL586973
c	29:	gb_est4:*	229	10.1	772	14	DX010757	KBrB001C0
c	30:	gb_est5:*	229	10.1	470	9	DN201839	USDA-FP_1
c	31:	gb_est6:*	225	10.0	254	7	BE520241	M1B12XTM
c	32:	gb_est7:*	219.5	9.7	692	5	CK564106	rswpb0_00
c	33:	gb_est8:*	192.5	8.5	551	5	CK535965	rswgb0_00
c	34:	gb_est9:*	185.5	8.2	687	10	DV079873	I13SFmidg
c	35:	gb_est0:*	185	8.2	692	9	CX700688	Hm_16w_04
c	36:	gb_est1:*	183.5	8.1	734	5	CK507686	rswda0_00
c	37:	gb_est3:*	182.5	8.1	575	5	CK543282	rswhb0_01
c	38:	gb_est4:*	177.5	7.9	696	1	AU004475	AU004475
c	39:	gb_est5:*	154	6.8	769	8	CO366885	AGENCOURT
c	40:	gb_est6:*	153	6.8	863	6	CV672453	RET7SJ_01
c	41:	gb_est7:*	147	6.5	1044	6	BX021319	Single_re
c	42:	gb_est8:*	144.5	6.4	645	2	BJ037937	BJ037937
c	43:	gb_est9:*	140	6.2	250	14	DX065699	KBrB07JU0
c	44:	gb_est3:*	132.5	5.9	2201	6	CR926080	Pongo_pyg
	45:	gb_est4:*	132	5.8	1430	6	CR615614	full-leng

ALIGNMENTS

RESULT 1
LOCUS   CL662734
LOCUS   CL662734   853 bp   DNA   linear   GSS 09-JUL-2004
DEFINITION   PRI0142b_B11 - PRI0142b.B21 (853) Mixed stage fosmid library of P.
               pacificus var. California Pristionchus pacificus genomic, genomic
               survey sequence.
ACCESSION   CL662734
VERSION   CL662734.1 GI:50150877
KEYWORDS   GSS.
SOURCE   Pristionchus pacificus
  ORGANISM   Pristionchus pacificus
               Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
               Neodiplogasteridae; Pristionchus.
REFERENCE   1 (bases 1 to 853)
  AUTHORS   Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
  TITLE   AppaDB: an AcedB database for the nematode satellite organism
               Pristionchus pacificus
  JOURNAL   Nucleic Acids Res. 32 (1), D421-D422 (2004)
  PUBMED   14681447
COMMENT   Contact: Sommer RJ
               Max-Planck-Institute for Developmental Biology
               Evolutionary Biology
               Spemannstr. 37-39, Tuebingen D-72076, Germany
               Tel: 00497071601371
               Fax: 00497071601498
               Email: ralf.sommer@tuebingen.mpg.de
               This library was generated at Caltech, Pasadena, USA and end
               sequenced at Vancouver, Canada.

Result No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1016	45.0	853	13	CL662734	CL662734 PRI0142b-
c 2	1001	44.3	616	2	BJ074127	BJ074127 BJ074127
3	905	40.1	529	3	BU713770	BU713770 SJAABUG02
4	888	39.3	746	4	CA093060	CA093060 SCCCL200

Seq primer: T7
Class: fosmid ends.

FEATURES   Location/Qualifiers
     source   1. .853
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Mixed stage fosmid library of P. pacificus
                     var. California"
                     /note="Vector: pEpifos-5 Fosmid vector"

ORIGIN

US-10-601-319-10 (1-432) x CL662734 (1-853)

Alignment Scores:
Pred. No.:
Score: 3.43e-92   Length: 853
Percent Similarity: 1016.00   Matches: 193
Best Local Similarity: 98.5%   Conservative: 0
Query Match: 45.0%   Mismatches: 3
DB: 13   Indels: 0
                                         Gaps: 0

Qy   237 SerMetLeuThrGluIlePheLeuLeuGlnGlnAlaLeuGlnGlyMetProGluProGlyTrp 256
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2   TCAATGCTGACGGAGATATTTCTTCCTGCAACAAGCACAGGGAAATGCCGGAGCCCGGGGTGG 61

Qy   257 GlyArgIleThrAspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPhe 276
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62   GGAAGGATCACCGATTCACACCAGTGGAACACTTTGCTAAGTTTGCATAACGGCAATTT 121

Qy   277 AspLeuLeuGlnArgThrProGluValAlaLaArgSerArgAlaThrProLeuLeuAspLeu 296
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   122 TATTTGCTACAACGCACGCCAGAGGTTGCCCGCCACCCCGCTTAATTGGATTTG 181

Qy   297 IleLysThrAlaLeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuPro 316
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   182 ATCATGGCAGCGGTTGACGCCCCATCACCGCAAAAACAGGCGTATGGTGTGACATTACCC 241

Qy   317 ThrSerValLeuPheIleAlaGlyHisHisAspThrAsnLeuAlaLeuAsnLeuCysGlyAlaLeu 336
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   242 ACTTCAGTACTGTTTAATTGCCGACACGATACTAATCTGGCAAATCTTCGGCGGCCACTG 301

Qy   337 GluLeuAsnTrpThrLeuArgProGlyTyrGlnProAspProAsnThrLeuLeuSerLeuVal 356
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   302 GAGCTCAAACTGGACGCTTCCCCGGTCAGCCGGATAAACAGCGCCCACGGTGGTGAACTGGTG 361

Qy   357 PheGluArgThrProArgArgLeuSerArgAspAsnSerGlnTrpIleGluGlnValSerLeuValPhe 376
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   362 TTTGAACGGCTCGGCGCTAAGCGGGATAACAGCCAGTGGATTGAGCAGGTTTCGCTCGTTC 421

Qy   377 GlnThrLeuGlnGlnMetArgGlyAspGlyTyrProLeuSerLeuAsnThrProProGlyGlu 396
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   422 CAGACTTTACAGCAGATGCGTGGCGACGGCTGTCATTAAATACGGCCGCCCGAGAG 481

Qy   397 ValLysGlyLeuThrLeuArgAlaGlyCysGlyLeuArgAsnAlaGlnGlyMetCysSerLeuAla 416
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   482 GTGAAAGGACTGACCCTGCGGGATGTGAAGAGCGGAAATGCGCAGGGCATGTGTTCGTTGGCA 541

Qy   417 GlyPheThrGlyIleValAsnGlnAlaAlaArgIleProAlaCysSerSerLeu 432
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   542 GGTTTTACGGGAAATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTG 589

---

us-10-601-319-10.rst   Page 2

REFERENCE
  AUTHORS   Kitayama,A., Teraaka,C., Mochii,M., Ueno,N., Shin-1,T. and
            Kohara,Y.
  TITLE   Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT   Contact: Tadasu Shin-1
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshin1@genes.nig.ac.jp
            The information of this clone is available through the following
            URL.
            http://xenopus.nibb.ac.jp.

FEATURES   Location/Qualifiers
     source   1. .616
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8355"
                     /clone="XL090j06"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 25"
                     /clone_lib="NIBB Mochii normalized Xenopus tailbud
                     library"

ORIGIN

US-10-601-319-10 (1-432) x BJ074127 (1-616)

Alignment Scores:
Pred. No.:
Score: 6.95e-91   Length: 616
Percent Similarity: 1001.00   Matches: 198
Best Local Similarity: 96.6%   Conservative: 0
Query Match: 44.3%   Mismatches: 7
DB: 2   Indels: 0
                                         Gaps: 0

Qy   89   ValAlaAlaAspGlyLeuLeuProLysCysGlyCysProGlnSerGlyGlnValAlaIleIle 108
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   615 GTRAGCCGACGGAATTGCTGCCGCAAAAAGGGCTGCCCCCGCAGTCTCGGTCAGGTCCGGATTATT 556

Qy   109 AlaAspValAspGluArgThrArgLysThrGlyGluAlaPheAlaAlaGlyLeuAlaPro 128
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   555 GCTGATGTCGACGAGCGTACCCGTAAAAACAGGCGAAGCCTTGCCGCCGGGTGGCGACT 496

Qy   129 AspCysAlaIleThrValHisThrGlnAlaAlaAspThrSerSerProAspProLeuPheAsn 148
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   495 GACTGTGCAATAACCGTACATACCCAGGCACAGATACGTCCAGTCCCCGATCCGTTATTTAAT 436

Qy   149 ProLeuLysThrGlyTyrValCysGlnLeuAspAsnAlaAsnValThrAspAlaIleLeuGlu 168
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   435 CCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCGAATGTGACTGACGCGATCCTCAGC 376

Qy   169 ArgIleAlaGlySerIleAlaAspPheThrGlyHisThrTyrGlnThrAlaPheArgGluLeu 188
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   375 AGGGCAGGAGGGTCAATTGCCGACTTTACCGGGCATCGGCAACGGGCGTTTCGCGAACTG 316

Qy   189 GluArgValLeuAsnPheProGlnSerAsnLeuCysLeuLysArgGluLysGlnAlaGlu 208
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   315 GAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGCCTTAAACGTGAGAAACAGGACGA 256

Qy   209 SerCysSerLeuThrGlnIleAlaLeuProSerGluLeuLysValSerAlaAspCysValSer 228
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   255 AGCTGTTCATTAACGCAGGGCATTACCATCGGAACTCAAGGTGAGCGCCGAACGGGTTCTCA 196

Qy   229 LeuThrGlyAlaValSerLeuAlaSerMetLeuThrGluIlePheLeuLeuGlnGlnAla 248
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   195 TTAACCGGTGCGGTGAAGCCTCGGCATCAATGCTGACGGAGATATTTCTCCTGCAACAAGCA 136

Qy   249 GlnGlyMetProGluProGlyTrpGlyArgIleThrAspSerHisGlnTrpAsnThrLeu 268

RESULT 2
BJ074127/c

LOCUS   BJ074127   616 bp   mRNA   linear   EST 29-SEP-2003
DEFINITION BJ074127 NIBB Mochii normalized Xenopus tailbud library Xenopus
            laevis cDNA clone XL090j06 5', mRNA sequence.
ACCESSION   BJ074127
VERSION   BJ074127.1 GI:17504316
KEYWORDS   EST.
SOURCE   Xenopus laevis (African clawed frog)
  ORGANISM Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Xenopus.
REFERENCE   1 (bases 1 to 616)

```
Db    135 CTGGGAATGCCGGAGCCCGGGGTGGGGAAGGATCACCGATTCACCAGTGGAACACCTTG 76
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    269 LeuSerLeuHisAsnAlaGlnPheAspLeuLeuGlnArgThrProGluValAlaArgSer 288

Db     75 CTAAGTTTGCATAACGGCCAATTTTATTTGCTACAACCGCACGCCAGAGGTTGCCGCAGC 16
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    289 ArgValAlaThrProLeu 293

Db     15 CCCGCCACCCGTTA 1
          ||||||||||||||
Qy    289 ArgAlaThrProLeu 293
```

```
RESULT 3
LOCUS            BU713770              529 bp    mRNA            linear    EST 23-OCT-2003
DEFINITION       SJAABUG02 Adult SjC 7/94 Schistosoma japonicum cDNA similar to
                 pdb|1DKP|A Chain A, Crystal Structure Of Phytase Complex Of
                 Escherichia Coli Phytase At Ph 6.6. Phytate Is Bound With Its
                 3-Phosphate In The Active Site. Hg2+ Cation Acts As An
                 Intermolecular Bridge, mRNA sequence.
ACCESSION        BU713770
VERSION          BU713770.1  GI:28321126
KEYWORDS         EST.
SOURCE           Schistosoma japonicum
  ORGANISM       Schistosoma japonicum
                 Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
                 Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE        1  (bases 1 to 529)
  AUTHORS        Hu,W., Yan,Q., Shen,D.K., Liu,F., Zhu,Z.D., Song,H.D., Xu,X.R.,
                 Wang,Z.J., Rong,Y.P., Zeng,L.C., Wu,J., Zhang,X., Wang,J.J.,
                 Xu,X.N., Wang,S.Y., Fu,G., Zhang,X.L., Wang,Z.Q., Brindley,P.J.,
                 McManus,D.P., Xue,C.L., Feng,Z., Chen,Z. and Han,Z.G.
  TITLE          Evolutionary and biomedical implications of a Schistosoma japonicum
                 complementary DNA resource
  JOURNAL        Nat. Genet. 35 (2), 139-147 (2003)
   PUBMED        12973349
COMMENT          Contact: Zeguang Han
                 Chinese National Human Genome Center at Shanghai
                 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
                 201203, P. R. China
                 Tel: 86-21-50801919(ex.45)
                 Fax: 86-21-50801922
                 Email: hanzg@chgc.sh.cn.
FEATURES             Location/Qualifiers
     source          1..529
                     /organism="Schistosoma japonicum"
                     /mol_type="mRNA"
                     /strain="Chinese (Anhui) strain"
                     /db_xref="taxon:6182"
                     /sex="Male and female"
                     /tissue_type="Whole body"
                     /dev_stage="Adult worms"
                     /lab_host="Mouse and rabbit"
                     /clone_lib="Adult SjC 7/94"
                     /note="Vector: Lambda ZAP-II XR.; Site_1: EcoR I; Site_2:
                     Xho I; Several hundred adult Schistosoma japonicum
                     (Anhui, P.R. China, strain), of mixed sex, were perfused
                     from the mesenteries of experimentally infected mice and
                     rabbits at the Queensland Institute of Medical Research,
                     Brisbane, Australia (QIMR), and stored for several months
                     in liquid nitrogen. Subsequently, mRNA was isolated at the
                     QIMR from lysates of these worms by oligo dT
                     chromatography, using a kit from Pharmacia. The mRNA was
                     then shipped to Clontech, Palo Alto, CA, USA, who
                     constructed a cDNA library. First strand synthesis was
                     primed with an oligo-dT-XhoI-primer and synthesized using
                     M-MLV reverse transcriptase. Second strand synthesis was
                     accomplished with RNAse H and T4 DNA polymerase. The
                     double stranded cDNA was ligated to EcoRI linkers,
                     digested with EcoRI and XhoI, and ligated into the
                     phagemid vector lambda ZAP II XR. After construction of
                     this directional library by Clontech, it was returned to
                     the QIMR. During analysis of the library at the QIMR, we
                     have found that a small percentage, 2% to 3%, of the
```

ORIGIN

    clones contain inserts that appear to be highly homologous
    to sequences from salmonid fishes, as determined by
    homology comparisons using BLAST and by Southern
    hybridization analysis to genomic DNA from salmon (Sigma
    Chemical Co., St. Louis, MO) under stringent washing
    conditions. The remainder of the clones appear to contain
    S. japonicum sequences."

US-10-601-319-10 (1-432) x BU713770 (1-529)

```
Alignment Scores:
Pred. No.:
Score:                3.12e-81
Score:                905.00            Length:                   529
Percent Similarity:   98.9%            Matches:                  174
Best Local Similarity:98.9%            Conservative:             0
Query Match:          40.1%            Mismatches:               2
DB:                   3                Indels:                   0
                                       Gaps:                     0
```

```
Qy    224 AlaAspCysValSerLeuThrThrGlyValAlaValSerLeuAlaSerMetLeuThrGluIlePhe 243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GCCGGACAATGTCTCATTAACACCGGTGCGGTGAGCCTTCGCATCAATGCTGACGGAGATATTT 60

Qy    244 LeuLeuGlnGlnAlaLeuGlyMetProGluProGlyTrpGlyArgIleThrAspSerHis 263
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTCCTGCAACAAGCACAGGGAATGCCGGAGCCCGGGGTGGGGAGGATCACCGATTCACAC 120

Qy    264 GlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGlnArgThrPro 283
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CAGTGGAACACCTTGCTAAGTTTGCATAACGGCCAATTTTATTTGCTACAACGGCACGCCA 180

Qy    284 GluValAlaArgSerArgValAlaThrProLeuLeuAspLeuIleIleThrAlaLeuThrPro 303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GAGGTTGCCGCGCCAGCCCGCCACCCGTTATTAGATTGATCAAGACAACAGGCTTGACGCCC 240

Qy    304 HisProProGlnLeuTyrGlyValThrLeuProThrSerValLeuProPheIleAla 323
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CATCCACCGCAAAAAACAGGCGTATGGGTGTGGACATTACCCACTTCAGTGCTGTTTATCGGC 300

Qy    324 GlyHisAspThrAsnLeuAlaAsnLeuGlyGlyValAlaLeuGluLeuAsnTrpThrLeuPro 343
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGACACGATACTAATTCTGGCAAATCTCGGCGGCGCCACTGGAGCTCAACTGGACCGCTTCC 360

Qy    344 GlyGlnProAspAlaAsnThrProProGlyGlyGlyLeuValLeuPheGluArgTrpArgArgLeu 363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGTCAGCCCGGATAACACCCGCCCCAGGTGGTGGACTGGTGTTGTTTGAACGCTGGCGTCGGCTA 420

Qy    364 SerAspAsnSerGlnIleTrpIleGlyValSerValPheGluIleThrLeuGlnIleMetArg 383
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AGCGGATAACAGACCCAGTGGATTCAGGTTTCCGTCTGTCTTCCAGACTTTACAGCAGATGGGT 480

Qy    384 AspIleLysThrProLeuSerLeuAsnThrProProGlyGlyValValLysLeu 399
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GATAAAACGCCCGCTGTCATTAAATACCGCCGCCCCGAGGAGGTGAAACTG 528
```

```
RESULT 4
LOCUS            CA093060              746 bp    mRNA            linear    EST 23-SEP-2003
DEFINITION       SCCCCL2001D10.b Cl2 Saccharum officinarum cDNA clone SCCCCL2001D10
                 3', mRNA sequence.
ACCESSION        CA093060
VERSION          CA093060.1  GI:34963673
KEYWORDS         EST.
SOURCE           Saccharum officinarum
  ORGANISM       Saccharum officinarum
                 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
                 clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
                 complex.
REFERENCE        1  (bases 1 to 746)
  AUTHORS        Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE          The libraries that made SUCEST
  JOURNAL        Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
```

COMMENT   Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 001 row: D column: 10
            Seq primer: M13/Forward primer.

FEATURES   Location/Qualifiers
     source   1..746
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCCCL2001D10"
                     /lab_host="XL1Blue MRF'"
                     /clone_lib="CL2"
                     /note="Organ: Pool of sugarcane calli submitted to low
                     (4oC) and high (37 C) temperature stress; Vector:
                     pBluerScript; Site_1: EcoRI; Site_2: XhoI; An
                     unidirectional cDNA library generated from [Pool of
                     sugarcane calli submitted to low (4oC) and high (37 C)
                     temperature stress]. cDNA was prepared from polyA+ mRNA
                     using ZAP - cDNA Synthesis Kit (Stratagene). The
                     double-strand cDNAs were fractionated in a sepharose CL-2B
                     40cm-columns and fragments sizing between 0.8 and 1.5 Kb
                     were directionally cloned into the vector. Details of
                     each source of RNA and library construction can be
                     obtained at http://sucest.lad.ic.unicamp.br/public"

ORIGIN

US-10-601-319-10 (1-432) x CA093060 (1-746)

Alignment Scores:
Pred. No.:   2.83e-79
Score: 888.00
Percent Similarity: 96.7%
Best Local Similarity: 96.7%
Query Match: 39.3%

Length: 746
Matches: 177
Conservative: 0
Mismatches: 6
Indels: 1
Gaps: 0

Qy   204 GlyLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeuLysValSer 223
          |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   199 GAGAAACAGGACAGGACCTTGCTAAGTTTGCATAAACGCGCAATTACCATCGGAACTCAAGGTGAGC 258

Qy   224 AlaAspProCysValSerLeuThrGlyAlaValSerMetLeuThrGlyIlePhe 243
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   259 GCCGACAATGTCATTAACCGGTGCGTAGCCTCGCATCAATGCTGACGGAGATATT 318

Qy   244 LeuLeuGlnGlnAlaGlnGlyMetProGlyTrpGlyArgIleThrAspSerHis 263
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   319 CTCCTCCAACAAGCACAGGGAATGCCCGAGCCCGGGGAAGGATCACCGGATTCACAC 378

Qy   264 GlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGlnArgThrPro 283
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   379 CAGTGGAACACGCTTGCTAAGTTTGCATAAACGCGGCAATTTATTTGCTACAACGCCACGGCCA 438

Qy   284 GluValAlaArgSerArgAlaThrProLeuLeuAspLeuLysThrAlaLeuThrPro 303
          |||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   439 GAGGTTGCCCGGCACCCGCCACCCCGTTATTAGATTTGATCAAGACAGCGTTGACGCCC 498

Qy   304 HisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeuPheIleAla 323
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   499 CATCCACCGCANAAACAGGCGTATGGTGTGACATTACCCACTTCACTGCTGTTTATCGGCC 558

Qy   324 GlyHisAspThrAsnLeuAsnLeuGlyAlaLeuGluLeuAsnTrpThrLeuPro 343
          ||||||||| ||||||| |||||||||||||||||||||||||||||||||||
Db   559 GGACACGGATACTAATCTGGCAAATCTCGGGCGCCGCCAAATCTCCGGACGCTTCC 618

Qy   344 GlyGlnProAspAsnThrProGlyGlyLeuValPheGluArgTrpArgArgLeu 363
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   618 ...

            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   619 GGTCAGCCCGGATAACACGCCCGCCAGGTGGGTGGAACTGGTTGTT-GAACCCTGGGGTCNGCTA 677
Qy   678 AGCGGATAACAGCCAGTCGGATTCAGGTTTCGCTGGTCTTCAGACTTTACAGACANAATGCGT 737

Qy   384 AspLysThr 386
          |||||||||
Db   738 GATAAAACG 746

RESULT 5
LOCUS   CF326092/c   595 bp   mRNA   linear   EST 18-AUG-2003
DEFINITION JMT1--05-B09.g1 AtJMT-overexpressing transgenic rice lambda phage
            cDNA library (JMT1) Oryza sativa (japonica cultivar-group) cDNA
            clone JMT1--05-B09, mRNA sequence.
ACCESSION   CF326092
VERSION   CF326092.1 GI:33800445
KEYWORDS   EST.
SOURCE   Oryza sativa (japonica cultivar-group)
ORGANISM   Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
            clade; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1 (bases 1 to 595)
AUTHORS   Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
            Song,S.I., Kim,J.K., Kim,Y.K., Nahm,B.H.
TITLE   Large-scale Sequencing Analysis of Rice ESTs
JOURNAL   Unpublished (2003)
COMMENT   Contact: Nahm B.H.
            Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
            of Bioscience and Bioinformatics, Myongji University
            Yongin, Kyeonggi, Korea
            Tel: 82 31 330 6193
            Fax: 82 31 321 6355
            Email: bhnahm@eggbio.com, bhnahm@bio.myongji.ac.kr.

FEATURES   Location/Qualifiers
     source   1..595
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nackdong"
                     /db_xref="taxon:39947"
                     /clone="JMT1--05-B09"
                     /tissue_type="leaf"
                     /dev_stage="14 days after germination"
                     /lab_host="E.coli SOLR"
                     /clone_lib="AtJMT-overexpressing transgenic rice lambda
                     phage cDNA library (JMT1)"
                     /note="Vector: pBluescript SK(+); Site 1: EcoRI; Site_2:
                     XhoI; cDNA was inserted into lamda Uni-ZAP XR vector at 5'
                     end with EcoRI and 3' end with XhoI site. mRNA was
                     prepared from Arabidopsis Jasmoinate Carboxyl
                     methyltransferase overexpression line."

ORIGIN

US-10-601-319-10 (1-432) x CF326092 (1-595)

Alignment Scores:
Pred. No.:   2.93e-75
Score: 847.00
Percent Similarity: 99.4%
Best Local Similarity: 99.4%
Query Match: 37.5%

Length: 595
Matches: 162
Conservative: 0
Mismatches: 1
Indels: 0
Gaps: 0

Qy   270 SerLeuHisAsnAlaGlnPheAspLeuLeuGlnArgThrProGluValAlaArgSerArg 289
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   595 AGTTTGCATAAACGCGCAATTTATTTGCTACAACGCCACGGCCAGAGGTTGCCCGGCACGC 536

Qy   290 AlaThrProLeuLeuAspLeuLysThrAlaLeuThrProHisProProGlnLysGln 309
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   535 GCCACCCCGTTATTAGATTTGATCAAGACAGCGTTGACGCCCCATCCACCGCAAAAACAG 476

RESULT 6

CN762997/c

LOCUS　　CN762997　　　895 bp　　mRNA　　linear　　EST 20-MAY-2004
DEFINITION　ID0AAA5DC09RM1 ApMS Acyrthosiphon pisum cDNA clone ID0AAA5DC09 5',
　　mRNA sequence.
ACCESSION　CN762997
VERSION　CN762997.1　GI:47536920
KEYWORDS　EST.
SOURCE　Acyrthosiphon pisum (pea aphid)
　ORGANISM　Acyrthosiphon pisum
　　Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
　　Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
　　Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
REFERENCE　1　(bases 1 to 895)
　AUTHORS　Hunter,W., Martinez-Torres,D., Rahbe,Y., Sabater-Munoz,B.,
　　Stern,D., Tagu,D. and Wincker,P.
　TITLE　An expressed sequence tags database for the pea aphid Acyrthosiphon
　　pisum
　JOURNAL　Unpublished (2004)
COMMENT　Contact: D. Tagu
　　INRA Rennes
　　UMR BiO3P, BP 35327, F-35653 Le Rheu Cedex France
　　Tel: +33.2.23.48.51.65
　　Fax: +33.2.23.48.51.50
　　Risk of contamination by bacterial sequences from obligatory
　　(Buchnera) or facultative endosymbionts. These sequences were
　　obtained in the frame of the International Consortium of Aphid
　　Genomics in collaboration with Genoscope
FEATURES　Location/Qualifiers
　source　1..895
　　/organism="Acyrthosiphon pisum"
　　/mol_type="mRNA"
　　/cultivar="developmentstage"
　　/db_xref="taxon:7029"
　　/clone="ID0AAA5DC09"
　　/tissue_type="whole insect"
　　/dev_stage="nymphs and adults (parthenogenetic females)"
　　/lab_host="XL1-Blue"
　　/note="Vector: pBS-SK minus; Site 1: EcoRI; Site 2: XhoI;
　　Sample name: ID0AAA ; Plant growth place: Department of
　　Ecology & Evolutionary Biology, Princeton University ;
　　Soil conditions: Soil ; Sowing date: 01/06/1999 ; Stress date: no stress ;
　　Harvesting date: 01/06/1999 ; Stress date: no stress ;
　　Description: Aphids inoculated on one-week old Vicia faba
　　under non-sterile conditions. All parthenogenetic stages
　　and both winged and wingless adults were collected for
　　library construction. / experimental condition: long
　　photoperiod (16-hr light/8-hr dark at 18 c)"

ORIGIN

US-10-601-319-10 (1-432) x CN762997 (1-895)

Alignment Scores:
Pred. No.:
Score:　　　3.57e-74　　　Length:　　　895
　　　　　839.00　　　Matches:　　　160
Percent Similarity:　99.4%　　Conservative:　0
Best Local Similarity:　99.4%　Mismatches:　1
Query Match:　　37.2%　　Indels:　　　0
DB:　　　　8　　　Gaps:　　　0

Qy　272　HisAsnAlaGlnPheAspLeuLeuGlnArgThrProGluValAlaArgSerArgAlaThr　291
　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　894　CATAACGCGCAATTTAATTTATTGCTACAACGCCACGCCGAGACGATACTAATCTGGCAAAT　835

Qy　292　ProLeuLeuAspLeuLeuIleLysThrAlaLeuThrProHisProProGlnLysGlnAlaTyr　311
　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　834　CCGTTATTAGATTTGATCAAGACAGCGTTGACGCCCCACCCCACAAAAACAGGCGTAT　775

Qy　312　GlyValThrLeuProThrSerValLeuPheIleAlaGlyHisAspThrAsnLeuAlaAsn　331
　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　774　GGTGTGACATTACCCACTTCAGTGCTGTTTATCGCCGGACACGATACTAATCTGGCAAAT　715

Qy　332　LeuGlyGlyAlaLeuGluLeuAsnTrpThrLeuProGlyProGlnProAspAsnThr　351
　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　714　CTCGGACGGGCGCTGGAGCTCAACTGGACAGCCTTCCCGGTCAACACGGATAACACGCGCCA　655

Qy　352　GlyGlyGlyLeuLeuValPheGluIleArgTrpArgArgLeuSerArgAspAspAsnSerGln　371
　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　654　GGTGGTGGAACTGGTGTTTGAACCGCTGGCGCTAAGCAGGGATGACAGCCAGTGGTTCAG　595

Qy　372　ValSerLeuValPheGlnThrLeuGlnMetArgArgAspLysThrProLeuSerLeuAsn　391
　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　594　GTTTCGCTGGTCTTCCAGACTTTACAGGATGCGGTGAGGATGTGAAGACGCCCGTGTCATTAAAT　535

Qy　392　ThrProProGlyProGlnValValLysGlyLeuThrLeuValAlaGlyLysAlaGlnGlyLy　411
　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　534　ACGCCGCGCCCGGAGAGGTGAAACTGGACCCTGGCAGGAGATGGCAGGAGGC　475

Qy　412　MetCysSerLeuValAlaGlyPheThrArgGlnIleValAlaAsnGluAlaArgIleProAlaCysSer　431
　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　474　ATGTGTTCGTTGGCCAGGTTTTACGGCAAATCGTGAATGAAGCACGCCAGT　415

Qy　432　Leu　432
　　|||
Db　414　TTG　412

RESULT 7

CN754382/c

LOCUS　　CN754382　　　868 bp　　mRNA　　linear　　EST 19-MAY-2004
DEFINITION　ID0AAA12DE01RM1 ApMS Acyrthosiphon pisum cDNA clone ID0AAA12DE01
　　5', mRNA sequence.
ACCESSION　CN754382
VERSION　CN754382.1　GI:47519379
KEYWORDS　EST.
SOURCE　Acyrthosiphon pisum (pea aphid)
　ORGANISM　Acyrthosiphon pisum
　　Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
　　Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
　　Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
REFERENCE　1　(bases 1 to 868)
　AUTHORS　Hunter,W., Martinez-Torres,D., Rahbe,Y., Sabater-Munoz,B.,
　　Stern,D., Tagu,D. and Wincker,P.
　TITLE　An expressed sequence tags database for the pea aphid Acyrthosiphon

JOURNAL   Unpublished (2004)
COMMENT   Contact: D. Tagu
           INRA Rennes
           UMR BiO3P, BP 35327, F-35653 Le Rheu Cedex France
           Tel: +33.2.23.48.51.65
           Fax: +33.2.23.48.51.50
           UMR BiO3P, BP 35327, F-35653 Le Rheu Cedex France
           Genomics in collaboration with Genoscope
           obtained in the frame of the International Consortium of Aphid
           (Buchnera) or facultative endosymbionts. These sequences were
           Risk of contamination by bacterial sequences from obligatory
           PCR PRimers
           FORWARD: CAGGAAACAGCTATGACC
           Plate: 12 row: E column: 1.
FEATURES   Location/Qualifiers
source   1. .868
           /organism="Acyrthosiphon pisum"
           /mol_type="mRNA"
           /cultivar="developmentstage"
           /db_xref="taxon:7029"
           /clone="ID0AAA12DE01"
           /tissue_type="whole insect"
           /dev_stage="nymphs and adults (parthenogenetic females)"
           /lab_host="Xli-Blue"
           /clone_lib="ApMS"
           /note="Vector: pBS-SK minus; Site 1: EcoRI; Site 2: XhoI;
           Sample name: ID0AAA ; Plant growth place: Department of
           Ecology & Evolutionary Biology, Princeton University ;
           Soil conditions: Soil ; Sowing date: 01/06/1999 ;
           Harvesting date: 01/06/1999 ; Stress date: no stress ;
           Description: Aphids inoculated on one-week old Vicia faba
           under non-sterile conditions. All parthenogenetic stages
           and both winged and wingless adults were collected for
           library construction. ; experimental condition: long
           photoperiod (16-hr light/8-hr dark at 18 c)"
ORIGIN

US-10-601-319-10 (1-432) x CN754382 (1-868)

Alignment Scores:
Pred. No.:
Score: 5.37e-68
                        778.00   Length: 868
                                      Matches: 151
Percent Similarity: 99.3%   Conservative: 0
Best Local Similarity: 99.3%   Mismatches: 1
Query Match: 34.5%   Indels: 1
DB: 8   Gaps: 0

```
Qy  281 ArgThrProGlyLeuValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLeuLysThrAla 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  866 CGCACGCCAGAGTT-GCCCGCAGCCCGCGGCCACCCCGGTTTATTAGATTTGATCAAGACAGCG 808

Qy  301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  807 TTGACCGCCCCANCCACCGCAAAAACAGGCGTNNGGTGTGACANTTACCCACTTCAGTGCTG 748

Qy  321 PheIleAlaGlyHisAspThrAsnLeuValAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  747 TTTATCGCCGGACACGATACTAATTCTGGCAAAATCTCGGCGGCCACTGGAGCTCAACTGG 688

Qy  341 ThrLeuProGlyGlnProAspAlaAsnThrProProGlyLeuGlyLeuValPheGluGlnArgTrp 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  687 ACGCTTCCCGGTCAGCCCGATAACACGCCGGTGGTGAACCTGGTTGTTTGAACGGCTGG 628

Qy  361 ArgArgLeuSerArgAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  627 CGTCGCGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 568

Qy  381 GlnMetArgGlyAspLysThrProLeuSerLeuAsnThrProProGlyGlyValValLysLeuThr 400
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  567 CAGATGCGTGGTGATAAAACGCCGCTGTCATTAAATACCGCCGCCGGAGGTGAAAACTGACC 508

Qy  401 LeuAlaGlyTyrCysGluGlyLeuArgAsnAlaLaGlnGlyMetCysSerSerLeuAlaGlyPheThrGln 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  507 CTGGCAGGATGTGAAGAGCGAAATGGGCAGGGCCATGTGTTCGTTGCGTTTGGCCAGGTTTTACGCCAA 448
```

RESULT 8
LOCUS   CN757565   864 bp   mRNA   linear   EST 20-MAY-2004
DEFINITION ID0AAA1DF09RM1 ApMS Acyrthosiphon pisum cDNA clone ID0AAA1DF09 5',
            mRNA sequence.
ACCESSION   CN757565
VERSION   CN757565.1 GI:47531488
KEYWORDS   EST.
SOURCE   Acyrthosiphon pisum (pea aphid)
ORGANISM   Acyrthosiphon pisum
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
            Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
REFERENCE   1 (bases 1 to 864)
AUTHORS   Hunter,W., Martinez-Torres,D., Rahbe,Y., Sabater-Munoz,B.,
            Stern,D., Tagu,D. and Wincker,P.
TITLE   An expressed sequence tags database for the pea aphid Acyrthosiphon
            pisum
JOURNAL   Unpublished (2004)
COMMENT   Contact: D. Tagu
            INRA Rennes
            UMR BiO3P, BP 35327, F-35653 Le Rheu Cedex France
            Tel: +33.2.23.48.51.65
            Fax: +33.2.23.48.51.50
            Risk of contamination by bacterial sequences from obligatory
            (Buchnera) or facultative endosymbionts. These sequences were
            obtained in the frame of the International Consortium of Aphid
            Genomics in collaboration with Genoscope
            PCR PRimers
            FORWARD: CAGGAAACAGCTATGACC
            Plate: 1 row: F column: 9.
FEATURES   Location/Qualifiers
source   1. .864
            /organism="Acyrthosiphon pisum"
            /mol_type="mRNA"
            /cultivar="developmentstage"
            /db_xref="taxon:7029"
            /clone="ID0AAA1DF09"
            /tissue_type="whole insect"
            /dev_stage="nymphs and adults (parthenogenetic females)"
            /lab_host="Xli-Blue"
            /clone_lib="ApMS"
            /note="Vector: pBS-SK minus; Site 1: EcoRI; Site 2: XhoI;
            Sample name: ID0AAA ; Plant growth place: Department of
            Ecology & Evolutionary Biology, Princeton University ;
            Soil conditions: Soil ; Sowing date: 01/06/1999 ; Stress date: no stress ;
            Harvesting date: 01/06/1999 ; Stress date: no stress ;
            Description: Aphids inoculated on one-week old Vicia faba
            under non-sterile conditions. All parthenogenetic stages
            and both winged and wingless adults were collected for
            library construction. ; experimental condition: long
            photoperiod (16-hr light/8-hr dark at 18 c)"
ORIGIN

US-10-601-319-10 (1-432) x CN757565 (1-864)

Alignment Scores:
Pred. No.:
Score: 5.53e-67
                        768.00   Length: 864
                                      Matches: 150
Percent Similarity: 99.3%   Conservative: 0
Best Local Similarity: 99.3%   Mismatches: 1
Query Match: 34.0%   Indels: 1
DB: 8   Gaps: 0

```
Qy  507 CTGGCAGGATGTGAAGAGCGAAATGGGCAGGGCCATGTGTTCGTTGCGTTTGGCCAGGTTTTACGCCAA 448
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 IleValAsnGluAlaArgIleProAlaCysSerSerLeu 432

Db  447 ATCGTGAATGAAGCACGCCATACCGGCCGCTGCCAGTTTG 412

Qy  282 ThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLeuLysThrAlaLeu 301
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  863 ACGCCAGAGGTTGCCCGCAGCCCGCGGCCACCCCGGTTTATTAGATTG-ATCAAGACAGCGTTG 805
```

```
Qy      302 ThrProHisProProGlyLysGlnAlaTyrGlyValThrLeuProThrSerValLeuPhe 321
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      804 ACGCCCCATCCACCGGAAAAACAGGCGTATAGGGNTGACATTACCCACTTCAGTGTCGTTT 745

Qy      322 IleAlaGlyHisAspThrAsnLeuLeuAlaAsnLeuGlyGlyValAlaLeuAsnTrpThr 341
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      744 ATCGCCGGACACGATACTAATCTGGCCAANTCCGGCGGGCCACTGGAGCTCAACTGGACG 685

Qy      342 LeuProGlyGlnProAspAsnThrProProGlyGlyGlyGluGluValPheGluArgTrpArg 361
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      684 CTTCCCCGGTCAGCCCGGATAACACGCCCCAGGTGGTGAACCTGGTGGGCT 625

Qy      362 ArgLeuSerAspAsnSerGlnThrProIleGlnIleThrLeuGlnGln 381
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db      624 CGGCTAACGGATAACAGCCAGTGGAATTCAGGTTTCGCTCGGTCTTCCAGACTTTACAGCAG 565

Qy      382 MetArgAspGlyThrProLeuSerLeuAlaAsnThrProProGlyGlyGluValLysLeuValLeuThrLeu 401
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      564 ATCCGTGATAAAAACGCCCGTCATTAAAATACCCGCCCCGGAGAGGGTGAAACTGACCCTG 505

Qy      402 AlaGlyLysCysGluGlyValAsnAlaGlnAsnLeuGlyMetCysSerLeuAlaLeuGlyPheThrGlnIle 421
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      504 GCAGGATGTGAAGAGCGAAAATGCCCAGCGGCCATGTGTTCCGTTCGGCAGGTTTTACGCAAATC 445

Qy      422 ValAsnGluValAlaArgIleProAlaCysSerLeu 432
            ||||||||||||||||||||||||||||||||||||
Db      444 GTGAATGAAGCACGCATACCGGCGTGCAGTTTG 412
```

```
RESULT 9
LOCUS       AW036132       383 bp    mRNA     linear     EST 18-MAY-2001
DEFINITION  EST274508 tomato seed, TAMU Lycopersicon esculentum cDNA clone
            cLEB1E23 similar to periplasmic phosphoanhydride phosphohydrolase
            precursor, putative, mRNA sequence.
ACCESSION   AW036132
VERSION     AW036132.1  GI:5894811
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum (Solanum lycopersicum)
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 383)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,
            Holt,I.E., Liang,F., Upton,J., Ronning,C.M., Fraser,C.M., Craven,M.B.,
            Fujii,C.Y., Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C.,
            Martin,G.B., Tanksley,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato seed tissue
  JOURNAL   Unpublished (1999)
  COMMENT   Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            3 prime sequence.
FEATURES            Location/Qualifiers
     source         1..383
                    /organism="Lycopersicon esculentum"
                    /mol_type="mRNA"
                    /cultivar="TA496"
                    /clone="cLEB1E23"
                    /db_xref="taxon:4081"
                    /tissue_type="seeds"
                    /dev_stage="quiescent seed"
                    /lab_host="Xl1-blue MRF'"
                    /clone_lib="tomato seed, TAMU"
                    /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
                    XhoI; cLEB - Tomato Seed EST Library. Directionally cloned
                    cDNAs inserted into pBlueScript SK(-) at 5' end with
                    EcoRI and 3' end with XhoI site."
ORIGIN
```

```
Alignment Scores:
Pred. No.:                           Length:         383
Score:          8.57e-53             Matches:        124
Percent Similarity:  623.00          Conservative:   0
Best Local Similarity: 97.6%         Mismatches:     3
Query Match:         27.6%           Indels:         0
DB:                  7               Gaps:           0
```

```
US-10-601-319-10 (1-432) x AW036132 (1-383)

Qy      117 LysThrGlyAlaAlaPheAlaAlaAlaGlyLeuValAlaProAspCysAlaAlaIleThrValHisSerThr 136
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2 AAAACAGGCGGAAGCCTTCGCCGCCGCCGGGCTGGCCACCTGACGTGCGCAATAACCGTACATACC 61

Qy      137 GlnAlaAspThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGln 156
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       62 CAGGCGGATACGTCCAGTCCCGATCCGTTGTTTAATCCCTTAAAAACTGGCGTTGTTGCCAA 121

Qy      157 LeuAspAsnAlaAsnValThrAspProAlaIleLeuGluAlaArgGlyGlySerIleAlaAsp 176
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      122 CTGGATAACGCGAACGTGACTGACCCGGCATCCTCAGCACGGCGGAGGGTCAATTGGCTGAC 181

Qy      177 PheThrGlyHisTyrGlnThrAlaPheArgGluGluGlyArgValLeuAsnPheProGln 196
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      182 TTTACCGGGCCATCGGCAAACGGCGTTTCCGGAACTGGAACGGGGTGCTTAATTTCCGCAA 241

Qy      197 SerAsnLeuCysSerLeuLysArgGluLysGlnSerCysSerSerLeuThrGlnAlaLeu 216
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      242 TCAAACTTGTGCCTTAAACGTGAGAAACAGGACGAAAGCGTGTTCATTAACGCAGCCAGTTA 301

Qy      217 ProSerGluLeuLysValSerArgAlaAspCysValSerArgLeuThrGlyAlaValSerLeuValAla 236
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      302 CCATCGGAACTCAAGGGTGAGGCCGCCGACAATGTCTCATTAACCGGTGCCGGTAAAGCCTCGCA 361

Qy      237 SerMetLeuThrGluIlePhe 243
            |||||||||||||||||||||
Db      362 TCAATGCTGACCGGAGATATTT 382
```

```
RESULT 10
LOCUS       AW036134/c       383 bp    mRNA     linear     EST 18-MAY-2001
DEFINITION  EST274510 tomato seed, TAMU Lycopersicon esculentum cDNA clone
            cLEB1E23 similar to periplasmic phosphoanhydride phosphohydrolase
            precursor, putative, mRNA sequence.
ACCESSION   AW036134
VERSION     AW036134.1  GI:5894813
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum (Solanum lycopersicum)
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 383)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,
            Holt,I.E., Liang,F., Upton,J., Ronning,C.M., Fraser,C.M., Craven,M.B.,
            Fujii,C.Y., Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C.,
            Martin,G.B., Tanksley,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato seed tissue
  JOURNAL   Unpublished (1999)
  COMMENT   Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES            Location/Qualifiers
     source         1..383
                    /organism="Lycopersicon esculentum"
                    /mol_type="mRNA"
                    /cultivar="TA496"
                    /clone="cLEB1E23"
                    /db_xref="taxon:4081"
                    /tissue_type="seeds"
```

/dev_stage="quiescent seed"
/lab_host="XL1-Blue MRF'"
/clone_lib="tomato seed, TAMU"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; clone="M11B12 5', mRNA sequence.
/note="Vector: pBluescript SK(-). Directionally cloned
cDNAs inserted into pBluescript SK(-) at 5' end with
EcoRI and 3' end with XhoI site."

ORIGIN

US-10-601-319-10 (1-432) x AW036134 (1-383)

Alignment Scores:
Pred. No.:
Score: 8.57e-53   Length: 383
Score: 623.00   Matches: 124
Percent Similarity: 97.6%   Conservative: 0
Best Local Similarity: 97.6%   Mismatches: 3
Query Match: 27.6%   Indels: 0
DB: 7   Gaps: 0

Qy 117 LysThrGlyGluAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThr 136
Db 382 AAAACAGGCGAAGCCTTCGCCGCGGGCTGCACCTGACTGTGCAATAACCGTACATACC 323

Qy 137 GlnAlaAspThrSerSerProAspPheAsnProLeuLysThrGlyValCysGln 156
Db 322 CAGGCAGATACGTCCAGTCCCGGATCCGTTATTTAATCCTCTAAAAAACTGGGGTTTCCAA 263

Qy 157 LeuAspAsnAlaAsnValThrAspAlaIleLeuLeuGluArgGlyAlaGlyLeuGlySerIleAlaAlaAsp 176
Db 262 CTGGATAACGCGGAACGTGACTGACGGCCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGAC 203

Qy 177 PheThrGlyHisTyrGlnThrThrAlaPheArgGluArgValLeuAsnPheProGln 196
Db 202 TTTACCGGGCATCGGCGAAACCGGGTTTCGGCGAACTGGAACCGGTGCTTAATTTCCGGCAA 143

Qy 197 SerAsnLeuCysSerLeuLysArgGlyLysGlnSerCysSerLeuThrGlnAlaLeu 216
Db 142 TCAAACTGTGTCCCTTAAACGTGGAAAACAGGACGAAAAAGCTGTTCATTAACGGAGCATTA 83

Qy 217 ProSerGluLeuLysValSerAlaAspCysValSerLeuThrHisGlyAlaValValSerLeuLeuAla 236
Db 82 CCATCCGGAACTCAAGGTGAGCCCGACAAATGTCTCATTAACCGGTGCGGTAAGCCTCGCA 23

Qy 237 SerMetLeuThrGluLeuPhe 243
Db 22 TCAATGCTGACGGAGATATTT 2

RESULT 11
LOCUS   BE520240/c
DEFINITION   M11B12STM Arabidopsis developing seed Arabidopsis thaliana cDNA
             clone M11B12 5', mRNA sequence.
ACCESSION   BE520240
VERSION   BE520240.1 GI:9778242
KEYWORDS   EST.
SOURCE   Arabidopsis thaliana
  ORGANISM   Arabidopsis thaliana (thale cress)
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
             rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1 (bases 1 to 393)
  AUTHORS   White,J.A., Todd,J., Newman,T., Focks,N., Girke,T., Martinez de
             Ilarduya,O., Jaworski,J.G., Ohlrogge,J. and Benning,C.
  TITLE   A new set of Arabidopsis expressed sequence tags from developing
             seeds. The metabolic pathway from carbohydrates to seed oil
  JOURNAL   Plant Physiol. 124 (4), 1582-1594 (2000)
   PUBMED   11115876
COMMENT   Contact: Benning, C
             Dept. of Biochemistry & Molecular Biology
             Michigan State University
             224 Biochemistry, Michigan State University, East Lansing, MI
             48824, USA
             Tel: 517 355 1609

Fax: 517 353 9334
Email: benning@msu.edu
Michigan State University DNA Sequencing Facility Arabidopsis
Biological Resource Center, The Ohio State University, 309 Botany &
Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX:
61429206 03 TEL: 614292971.

FEATURES   Location/Qualifiers
     source   1..393
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /ecotype="Columbia"
                     /db_xref="taxon:3702"
                     /clone="M11B12"
                     /tissue_type="seed"
                     /dev_stage="5-13 days after flowering"
                     /lab_host="E.coli"
                     /clone_lib="Arabidopsis developing seed"
                     /note="Organ: Developing seed; Vector: pBluescript SK-;
                     Site_1: EcoRI; Site_2: XhoII"

ORIGIN

US-10-601-319-10 (1-432) x BE520240 (1-393)

Alignment Scores:
Pred. No.:   1   Length: 393
Score: 1.42e-52   Matches: 124
Score: 621.00   Conservative: 1
Percent Similarity: 95.4%   Mismatches: 6
Best Local Similarity: 94.7%   Indels: 0
Query Match: 27.5%   Gaps: 0
DB: 7

Qy 107 IleIleAlaAspValAsnAspGluArgThrArgLysThrGlyGluAlaPheAlaAlaGlyLeu 126
Db 393 ATTATTGCTGATGTCGACGACGGTACCCGTAANACAGGCGAAGCTTTCGCGGCAGGGNTG 334

Qy 127 AlaProAspCysAlaIleThrValHisThrGlnAlaLeuProSerGluLeu 146
Db 333 GCACCTGACTGTGCAATAACCGTACATACCCAGGCAGAITACGTCCGGATCCGTTA 274

Qy 147 PheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAlaAsnValThrAspAlaIle 166
Db 273 TTTAATCCTCTAAAAAACTGGGGTTTGCCAACTGGATAACGCGAACGTGACTGACGGATC 214

Qy 167 LeuGluArgAlaLaGlyLyGlySerIleAlaAlaAspPheThrGlyHisTyrGlnThrThrAlaPheArg 186
Db 213 CTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCATCGGCGAAACCGGGTTTCGC 154

Qy 187 GluLeuValArgValLeuAsnPheProGlnSerAsnLeuCysSerLeuLysArgGlyLysGln 206
Db 153 GAACTGGAACCGGTGCTTAATTTCCGCCAATCAAACTGTGTCCCTTAAACGTGAGAAACAG 94

Qy 207 AspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeuLysValSerAlaAspCys 226
Db 93 GACGAAAAGCTGTTCATTAACCGGCCATTAACCATCCGGAACTCAAGGTGAGCCGCCGACAAT 34

Qy 227 ValSerLeuThrGlyAlaValValSerLeuLeuAlaSer 237
Db 33 GTCTCAITTAACCGCTCGCGGTAAGCCTCGCATCA 1

RESULT 12
LOCUS   BU713769   531 bp   mRNA   linear   EST 23-OCT-2003
DEFINITION   SJAABUG01, Adult SjC 7/94 Schistosoma Japonicum cDNA similar to
             spl P07102|PPA_ECOLI PERIPLASMIC APPA PROTEIN PRECURSOR [INCLUDES:
             PHOSPHOANHYDRIDE PHOSPHOHYDROLASE (PH 2.5 ACID PHOSPHATASE) (AP);
             6-PHYTASE], mRNA sequence.
ACCESSION   BU713769
VERSION   BU713769.1 GI:28321125
KEYWORDS   EST.
SOURCE   Schistosoma japonicum
  ORGANISM   Schistosoma japonicum
             Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
             Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.

```
REFERENCE   1 (bases 1 to 531)
AUTHORS     Hu,W., Yan,Q., Shen,D.K., Liu,F., Zhu,Z.D., Song,H.D., Xu,X.R.,
            Wang,Z.J., Rong,Y.P., Zeng,L.C., Wu,J., Zhang,X., Wang,J.J.,
            Xu,X.N., Wang,S.Y., Fu,G., Zhang,X.L., Wang,Z.Q., Brindley,P.J.,
            McManus,D.P., Xue,C.L., Feng,Z., Chen,Z. and Han,Z.G.
TITLE       Evolutionary and biomedical implications of a Schistosoma japonicum
            complementary DNA resource
JOURNAL     Nat. Genet. 35 (2), 139-147 (2003)
PUBMED      12973349
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn.
FEATURES             Location/Qualifiers
     source          1..531
                     /organism="Schistosoma japonicum"
                     /mol_type="mRNA"
                     /strain="Chinese (Anhui) strain"
                     /db_xref="taxon:6182"
                     /sex="Male and female"
                     /tissue_type="whole body"
                     /dev_stage="Adult worms"
                     /lab_host="Mouse and rabbit"
                     /clone_lib="Adult SjC 7/94"
                     /note="Vector: Lambda ZAP-II XR.; Site_1: EcoR I; Site_2:
                     Xho I; Several hundred adult Schistosoma japonicum
                     (Anhui, P.R. China, strain), of mixed sex, were perfused
                     from the mesenteries of experimentally infected mice and
                     rabbits at the Queensland Institute of Medical Research,
                     Brisbane, Australia (QIMR), and stored for several months
                     in liquid nitrogen. Subsequently, mRNA was isolated at the
                     QIMR from lysates of these worms by oligo dT
                     chromatography, using a kit from Pharmacia. The mRNA was
                     then shipped to Clontech, Palo Alto, CA, USA, who
                     constructed a cDNA library. First strand synthesis was
                     primed with an oligo-dT-XhoI-primer and synthesized using
                     M-MLV reverse transcriptase. Second strand synthesis was
                     accomplished with RNAse H and T4 DNA polymerase. The
                     double stranded cDNA was ligated to EcoRI linkers,
                     digested with EcoRI and XhoI, and ligated into the
                     phagemid vector lambda ZAP II XR. After construction of
                     this directional library by Clontech, it was returned to
                     the QIMR. During analysis of the library at the QIMR, we
                     have found that a small percentage, 2% to 3%, of the
                     clones contain inserts that appear to be highly homologous
                     to sequences from salmonoid fishes, as determined by
                     homology comparisons using BLAST and by Southern
                     hybridization analysis to genomic DNA from salmon (Sigma
                     Chemical Co., St. Louis, MO) under stringent washing
                     conditions. The remainder of the clones appear to contain
                     S. japonicum sequences."
```

```
ORIGIN

US-10-601-319-10 (1-432) x BU713769 (1-531)

Alignment Scores:
Pred. No.:
Score:                  1.46e-48
                        583.50
Percent Similarity:     79.5%          Length:        531
Best Local Similarity:  75.6%          Matches:       133
Query Match:            25.8%          Conservative:  7
DB:                     3              Mismatches:    30
                                       Indels:        7
                                       Gaps:          1

Qy  224 AlaAspCysValSerLeuThrGlyAlaAlaValSerLeuValAlaSerMetLeuThrGluIlePhe 243
Db    1 GCCGACAATGTCTCATTAACCGGTGCCGTAAGCCTCGCATCAATGCTGACGGAGATGGGG 60

Qy  244 LeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrTrpGlyArgIleThrAspSerHis 263
Db   61 CTCCTGCAACAAGCACAGGGAATGCCCGAGCCGGGGTGGGGAGGAGGATCACCGATTCACAC 120
```

```
RESULT 13
CN759004/c
LOCUS       CN759004               706 bp    mRNA    linear   EST 20-MAY-2004
DEFINITION  ID0AAA24BC04RM1 ApMS Acyrthosiphon pisum cDNA clone ID0AAA24BC04
            5', mRNA sequence.
ACCESSION   CN759004
VERSION     CN759004.1  GI:47532927
KEYWORDS    EST.
SOURCE      Acyrthosiphon pisum (pea aphid)
  ORGANISM  Acyrthosiphon pisum
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
            Aphidoidea; Aphididae; Aphidinae; Macrosiphini; Acyrthosiphon.
REFERENCE   1 (bases 1 to 706)
AUTHORS     Hunter,W., Martinez-Torres,D., Rahbe,Y., Sabater-Munoz,B.,
            Stern,D., Tagu,D. and Wincker,P.
TITLE       An expressed sequence tags database for the pea aphid Acyrthosiphon
            pisum
JOURNAL     Unpublished (2004)
COMMENT     Contact: D. Tagu
            Unpublished
            INRA Rennes
            UMR BiO3P, BP 35327, F-35653 Le Rheu Cedex France
            Tel: +33.2.23.48.51.65
            Fax: +33.2.23.48.51.50
            Risk of contamination by bacterial sequences from obligatory
            (Buchnera) or facultative endosymbionts. These sequences were
            obtained in the frame of the International Consortium of Aphid
            Genomics in collaboration with Genoscope
            PCR Primers
            FORWARD: CAGGAAACAGCTATGACC
            Plate: 24  row: C  column: 4.
FEATURES             Location/Qualifiers
     source          1..706
                     /organism="Acyrthosiphon pisum"
                     /mol_type="mRNA"
                     /cultivar="developmentstage"
                     /db_xref="taxon:7029"
                     /clone="ID0AAA24BC04"
                     /tissue_type="whole insect"
                     /dev_stage="nymphs and adults (parthenogenetic females)"
                     /lab_host="xll-Blue"
                     /clone_lib="ApMS"
                     /note="Vector: pBS-SK minus; Site_1: EcoRI; Site_2: XhoI;
                     Sample name: ID0AAA ; Plant growth place: Department of
```

```
Qy  264 GlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnProPheAspLeuLeuGlnArgThrPro 283
Db  121 CAGTGGAACACCTTGCTAAGTTTGCATAAGGCCCAATTTTATTTGCTACAACGCACGGCCA 180

Qy  284 GluValAlaArgSerArgAlaThrProLeuLeuAspLeuValLysThrAlaLeuThrPro 303
Db  181 GAGGTTGCCCGCAGCCGCGGCCACCCCGGGTATTAGATTTGACAGACAGGGTTGACGCCC 240

Qy  304 HisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu-PheIleAl 323
Db  241 CGTTTATCATTATGATTTGCATGTGGAGACAATTCCCTTGAGTGACGAAAAAAGG 300

Qy  323 aGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrpThrLeuPr 343
Db  301 CGGGCCACGACGAGCCTGAGCCTGATAAAAACTCGGCGGCGGGGGGCTCAACTGGACGCC 360

Qy  343 oGlyGlnProAspAspAsnThrProProGlyGlyGluLeu-ValPheGluAspTrp-ArgArg 362
Db  361 CAGGGACAGCCCGATAACACGCCCCGAGGAGGGGAACTGGATGTTTGAACGCGGTCGG 420

Qy  363 LeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGlnGlnMet 382
Db  421 CTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTA-CAGGAGATG 479

Qy  383 ArgAspLysThrProLeuSerLeuAsnThrProProGlyGlu 396
Db  480 CGTCGGGGGGGGTTCCTTTGGGG-----------CCTTATGGGGAT 512
```

ORIGIN
Ecology & Evolutionary Biology, Princeton University ;
Soil conditions: Soil ; Sowing date: 01/06/1999 ;
Harvesting date: 01/06/1999 ; Stress date: no stress ;
Description: Aphids inoculated on one-week old Vicia faba
under non-sterile conditions. All parthenogenetic stages
and both winged and wingless adults were collected for
library construction.; experimental condition: long
photoperiod (16-hr light/8-hr dark at 18 c)"
/clone="HBMSFIB4"
/sex="Male and Female"
/tissue_type="bone marrow"
/cell_type="stromal fibroblast"
/dev_stage="mixed"
/lab_host="XL1-Blue"
/clone_lib="Human Bone Marrow Stromal Fibroblast"
/note="Vector: pBluescript; Site_1: EcoRI; Site_2: XhoI"

US-10-601-319-10 (1-432) x CN759004 (1-706)

Alignment Scores:
Pred. No.:   4.02e-42
Score: 522.00
Percent Similarity: 100.0%   Length: 706
Best Local Similarity: 100.0%   Matches: 98
Query Match: 23.1%   Conservative: 0
DB: 8   Mismatches: 0
                                Indels: 0
                                Gaps: 0

Qy 335 AlaLeuGluLeuAsnTrpThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlu 354
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 705 GCACTGGAGCTCAACTGGACCCTTCCCGGTCAGCCCGATAACACACGGCCCGGTGGTGAA 646

Qy 355 LeuValPheGluArgThrArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeu 374
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 645 CTGGTGTTTGAACGCACGCGGCGTCGGCTAAGCGGATAACAGCCAGTGGATTTCAGGTTTCGCTG 586

Qy 375 ValPheGlnThrArgLeuGlnHisMetArgArgAspLysThrProLeuSerLeuAsnThrProPro 394
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 585 GTCTTCCAGACTTTACAGCAGATGCCGTGATAAAACCCCGCGTGTCATTAAAAACGGCCGCCCC 526

Qy 395 GlyGluValLysLeuThrLeuAlaGlyCysGluGlyValArgAsnAlaGlnGlyMetCysSer 414
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 525 GGAGAGGTGAAAACTGACCCTGCCAGGATGTGAAGAGGCGAAATGCCGAGGGCATGTGTTCG 466

Qy 415 LeuAlaGlyPheThrGlnIleIleValAsnGluLaAlaArgIleProAlaCysSerLeu 432
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 465 TTGGCAGGTTTTACGCAAAATCGTGAATGAACGCACGCATACCCGCGGTGCCAGTTTG 412

RESULT 14
AA545747/c
LOCUS   AA545747   354 bp   mRNA   linear   EST 12-MAY-1999
DEFINITION   HBMSFIB4-REV Human Bone Marrow Stromal Fibroblast Homo sapiens cDNA
             clone HBMSFIB4 5', mRNA sequence.
ACCESSION   AA545747
VERSION   AA545747.1 GI:2307026
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Hominidae; Homo.
REFERENCE   1 (bases 1 to 354)
AUTHORS   Jia,L., Robey,P., Young,M., Schuler,G., Powell,J., Yang,L.,
             Lennon,G., Hillier,L., Allen,M., Bowles,L., Geisel,S., Kucaba,T.,
             Marra,M., Martin,J., Steptoe,M., Wilson,R. and Francomano,C.
             Wylie,T., Waterston,R.,
TITLE   WashU-MGB/NHGRI EST Project
JOURNAL   Unpublished (1997)
COMMENT   Contact: Libin Jia
             Medical Genetics Branch
             National Human Genome Research Institute
             10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
             Tel: 301-402-4877
             Fax: 301-496-7157
             Email: libin@helix.nih.gov
             Seq primer: M13 Reverse.
FEATURES   Location/Qualifiers
     source   1..354
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"

US-10-601-319-10 (1-432) x AA545747 (1-354)

Alignment Scores:
Pred. No.:   1.28e-38
Score: 483.00
Percent Similarity: 91.6%   Length: 354
Best Local Similarity: 90.8%   Matches: 108
Query Match: 21.4%   Conservative: 1
DB: 1   Mismatches: 9
                                Indels: 4
                                Gaps: 0

Qy 166 IleLeuLeuGluArgAlaGlyGlySerIleAlaAlaSpPheThrGlyHisTyrGlnThr-AlaPh 185
        ::: ||||||||||||||||||||
Db 352 GTCTTCAGCAGGCGAGGAGGTCA-ATTGCTGA-TTTACCGCATCGGCAAACGGGGCGGTT 295

Qy 185 eArgGluLeuGlnArgValLeuAsnPheProGlnSerAsnLeuCysLeuLeuArgGluGly 205
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 294 TCGCGGAACTGGAACGGGTGTT-AATTTTCCGCAATCAAACTTGTGCCTTAAACGTGAGA 236

Qy 205 sGlnArgGluSerCysSerLeuThrGlnHisAlaLeuProSerGluLeuLysValSerAlaAs 225
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 235 ACAGGACGAGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTTCAAGGTTGCGCCGA 176

Qy 225 pCysValSerLeuThrGlyTyrAlaValSerLeuAlaSerMetLeuThrGluIlePheLeuLeu 245
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 175 CAATGTNTCATTAACCGGTGCGCGTAAGCCTCCCATCAATGCTGAGGAGATTATTTNTCCT 116

Qy 245 uGlnGlnAlaGlnGlyMetProGluProGlyTyrTrpGlyArgIleThrAspSerHisGlnTr 265
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 115 GCAACAAGCACAGGGAATGCCGAGCCCGGGTGGGGAAGGGATCACCGATTCACCAGTG 56

Qy 265 pAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGlnArgThrPro 283
        ::::::::::::   :::::::::
Db 55 GAACACCTTGCTAAGTTTGCATAACCGCAAATTTAATTGCTACAACGCACACCCA 1

source   1. .707
                /organism="Xenopus laevis"
                /mol_type="mRNA"
                /db_xref="taxon:8355"
                /clone="XL189d24"
                /tissue_type="whole embryo"
                /dev_stage="stage 10.5"
                /clone_lib="NIBB Mochii normalized Xenopus early gastrula
                library"

ORIGIN

Alignment Scores:
Pred. No.:
Score: 4.29e-34   Length: 707
Percent Similarity: 443.00   Matches: 84
Best Local Similarity: 98.8%   Conservative: 0
Query Match: 98.8%   Mismatches: 1
DB: 19.6%   Indels: 0
                        2   Gaps: 0

US-10-601-319-10 (1-432) x BJ619443 (1-707)

Qy   251 MetProGluProGlyTyrProGlyTrpGlyArgIleThrAspSerHisGlnTrpAsnThrLeuLeuSer 270
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2 ATGCCGGAGCCCGGGTGGGGAAGGATCACCGATTCACACCAGTGGGAACACCTTGCTAAGT   61

Qy   271 LeuHisAsnAlaGlnPheAspLeuLeuGlnArgThrProGluValAlaAlaArgSerArgAla 290
           |||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   62 TTGCATAACGCGCAATTTATTTGCTACAACGCACGCCAGAGGTTGCCCGCAGCGCGCC   121

Qy   291 ThrProLeuLeuAspLeuIleLysThrAlaLeuThrProHisProProGlnLysGlnAla 310
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   122 ACCCCGTTATTAGATTTGATCAAGACAGCGTTGAGCGCCCATCCACCGCAAAAACAGGCG 181

Qy   311 TyrGlyValThrLeuProThrSerValLeuPheIleAlaGlyHisAspThrAsnLeuAla 330
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   182 TATGGTGTGACATTACCCACTTCAGTGGTTATCGGCGGACACGGATACTAATCTGGCA 241

Qy   331 AsnLeuGlyGlyAla 335
           |||||||||||||||
Db   242 AATCTCGGCGGCGCA 256

Search completed: June 14, 2006, 15:26:34
Job time : 4899.67 secs

THIS PAGE LEFT BLANK

OM protein - protein search, using sw model

Copyright (c) 1993 - 2006 Biocceleration Ltd.

GenCore version 5.1.9

Run on: June 13, 2006, 10:17:34 ; Search time 189.248 Seconds
(without alignments)
2111.554 Million cell updates/sec

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_7.2:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Title: US-10-601-319-10
Perfect score: 2258
Sequence: 1 MKAILIPFLSLLIPLTPQSA..........CSLAGPTQIVNEARIPACSL 432

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2182	96.6	432	1	PPA_ECOLI	P07102 escherichia
2	2182	96.6	432	2	Q323E1_SHISS	Q323E1 shigella so
3	2178	96.5	432	2	Q8GN88_ECOLI	Q8gn88 escherichia
4	2176	96.4	432	2	Q8RKD7_ECOLI	Q8rkd7 escherichia
5	2176	96.4	432	2	Q8RKD7_ECOLI	Q8rkd7 escherichia
6	2176	96.4	432	2	Q8RKD8_ECOLI	Q8rkd8 escherichia
7	2176	96.4	432	2	Q8RKE0_ECOLI	Q8rke0 escherichia
8	2175	96.3	432	2	Q8RKD9_ECOLI	Q8rkd9 escherichia
9	2172	96.2	432	2	Q8RKE0_ECOLI	Q8rke0 escherichia
10	2172	96.2	432	2	Q8RKD5_ECOLI	Q8rkd5 escherichia
11	2154	95.4	432	2	Q31YP3_SHIBS	Q31yp3 shigella bo
12	2149	95.2	432	2	Q83RW2_SHIFL	Q83rw2 shigella fl
13	2147	95.1	432	2	Q6RK08_ECOLI	Q6rk08 escherichia
14	2145	95.0	432	2	Q7UD08_SHIFL	Q7ud08 shigella fl
15	2144	95.0	434	2	Q7AFW3_ECO57	Q7afw3 escherichia
16	2144	95.0	434	2	Q8XC29_ECO57	Q8xc29 escherichia
17	2143	94.9	446	2	Q8CW75_ECOL6	Q8cw75 escherichia
18	2130	94.3	434	2	Q32HS7_SHIDS	Q32hs7 shigella dy
19	1370.5	60.7	433	2	Q676V7_CITFR	Q676v7 citrobacter
20	1357.5	60.1	433	2	Q2VY22_CITBR	Q2vy22 citrobacter
21	1087	48.1	444	2	Q6U677_9ENTR	Q6u677 obesumbacte
22	1063.5	47.1	444	2	Q6TAQ8_9ENTR	Q6taq8 obesumbacte
23	959.5	42.5	441	2	Q669R3_YERPS	Q669r3 yersinia ps
24	959.5	42.5	441	2	Q8ZP6_YERPE	Q8zfp6 yersinia pe
25	560	24.8	413	2	Q8XBZ6_ECO57	Q8xbz6 escherichia
26	556	24.6	413	2	Q323B9_SHISS	Q323b9 shigella so
27	554	24.5	413	1	AGP_ECOLI	P19926 escherichia
28	553	24.5	413	2	Q5PG89_SALPA	Q5pg89 salmonella
29	553	24.5	413	2	Q83RV6_SHIFL	Q83rv6 shigella fl
30	552	24.4	413	1	AGP_SALTY	P33921 salmonella
31	547	24.2	413	2	Q32HQ8_SHIDS	Q32hq8 shigella dy
32	547	24.2	413	2	Q8Z7P1_SALTI	Q8z7p1 salmonella
33	538	23.8	417	1	AGP_PRORE	Q52309 providencia
34	536	23.7	413	2	Q6EV19_ENTCL	Q6ev19 enterobacte
35	536	23.7	414	2	Q9AAQ4_CAUCR	Q9aaq4 caulobacter
36	515	22.8	392	2	Q8PP76_XANAC	Q8pp76 xanthomonas
37	512	22.7	443	2	Q4UR06_XANC8	Q4ur06 xanthomonas
38	512	22.7	443	2	Q8P330_XANCP	Q8p330 xanthomonas
39	491	21.7	515	2	Q3BXB8_XANC5	Q3bxb8 xanthomonas
40	485	21.5	532	2	Q5GW75_XANOR	Q5gw75 xanthomonas
41	477.5	21.1	435	2	Q6CZF4_ERWCT	Q6czf4 erwinia car
42	475.5	21.1	435	2	Q8PF53_XANAC	Q8pf53 xanthomonas
43	471.5	20.9	433	2	Q3BMQ7_XANC5	Q3bmq7 xanthomonas
44	446	19.8	428	2	Q8GD20_PSESX	Q8gd20 pseudomonas
45	436	19.3	368	2	Q2NZE8_XANOR	Q2nze8 xanthomonas

ALIGNMENTS

RESULT 1
PPA_ECOLI

ID	PPA_ECOLI	STANDARD;	PRT;	432 AA.

AC P07102;
DT 01-APR-1988, integrated into UniProtKB/Swiss-Prot.
DT 01-AUG-1991, sequence version 2.
DT 07-MAR-2006, entry version 63.
DE Periplasmic appA protein precursor [Includes: Phosphoanhydride
DE phosphohydrolase (EC 3.1.3.2) (pH 2.5 acid phosphatase) (AP); 4-
DE phytase (EC 3.1.3.26)].
GN Name=appA; OrderedLocusNames=b0980;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND PROTEIN SEQUENCE OF 23-29.
RC STRAIN=K12;
RX MEDLINE=90368616; PubMed=2168385;
RA Dassa J., Marck C., Boquet P.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase.";
RL J. Bacteriol. 172:5497-5500(1990).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12 / W3110;
RX MEDLINE=97061202; PubMed=8905232; DOI=10.1093/dnares/3.3.137;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
RN [4]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 1-112.
RX MEDLINE=87271766; PubMed=3038201; DOI=10.1016/0300-9084(87)90045-9;
RA Touati E., Danchin A.;
RT "The structure of the promoter and amino terminal region of the pH 2.5
RT acid phosphatase structural gene (appA) of E. coli: a negative control
RT of transcription mediated by cyclic AMP.";

```
RL   Biochimie 69:215-221(1987).
RN   [5]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 1-17.
RC   STRAIN=K12;
RX   MEDLINE=92049231; PubMed=1658595; DOI=10.1007/BF00267454;
RA   Dassa J., Fsihi H., Marck C., Dion M., Kiefer-Bontemps M.,
RA   Boquet P.L.;
RT   "A new oxygen-regulated operon in Escherichia coli comprises the genes
RT   for a putative third cytochrome oxidase and for pH 2.5 acid
RT   phosphatase (appA).";
RL   Mol. Gen. Genet. 229:341-352(1991).
RN   [6]
RP   CHARACTERIZATION, AND PROTEIN SEQUENCE OF 23-34.
RA   Greiner R., Jany K.-D.;
RT   "Characterization of a phytase from Escherichia coli.";
RL   Biol. Chem. Hoppe-Seyler 372:664-665(1991).
RN   [7]
RP   CHARACTERIZATION, AND PROTEIN SEQUENCE OF 23-35.
RX   MEDLINE=93255556; PubMed=8387749; DOI=10.1006/abbi.1993.1261;
RA   Greiner R., Konietzny U., Jany K.-D.;
RT   "Purification and characterization of two phytases from Escherichia
RT   coli.";
RL   Arch. Biochem. Biophys. 303:107-113(1993).
RN   [8]
RP   CHARACTERIZATION.
RA   Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
RA   van Etten R.L.;
RT   "Overexpression, site-directed mutagenesis, and mechanism of
RT   Escherichia coli acid phosphatase.";
RL   J. Biol. Chem. 267:22830-22836(1992).
RN   [9]
RP   MUTAGENESIS.
RX   MEDLINE=93054596; PubMed=1429631;
RA   Golovan S., Wang G., Zhang J., Forsberg C.W.;
RT   "Characterization and overproduction of the Escherichia coli appA
RT   encoded bifunctional enzyme that exhibits both phytase and acid
RT   phosphatase activities.";
RL   Can. J. Microbiol. 46:59-71(2000).
RN   [10]
RP   X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS).
RX   MEDLINE=20122624; PubMed=10655611; DOI=10.1038/72371;
RA   Lim D., Golovan S., Forsberg C.W., Jia Z.;
RT   "Crystal structures of Escherichia coli phytase and its complex with
RT   phytate.";
RL   Nat. Struct. Biol. 7:108-113(2000).
CC   -!- CATALYTIC ACTIVITY: A phosphate monoester + H(2)O = an alcohol +
CC       phosphate.
CC   -!- CATALYTIC ACTIVITY: Myo-inositol hexakisphosphate + H(2)O = 1D-
CC       myo-inositol 1,2,3,4,5-pentakisphosphate + phosphate.
CC   -!- SUBUNIT: Monomer.
CC   -!- SUBCELLULAR LOCATION: Periplasmic.
CC   -!- INDUCTION: In addition to cAMP-mediated control, this enzyme is
CC       induced when bacterial cultures reach stationary phase; its
CC       synthesis is triggered by phosphate starvation or a shift from
CC       aerobic to anaerobic conditions.
CC   -!- SIMILARITY: Belongs to the histidine acid phosphatase family.
```

```
DR   EMBL; M58708; AAA72086.1; -; Genomic_DNA.
DR   EMBL; U00096; AAC74065.1; -; Genomic_DNA.
DR   EMBL; D90735; BAA35745.1; -; Genomic_DNA.
DR   EMBL; X05471; CAA29031.1; -; Genomic_DNA.
DR   EMBL; M58708; -; NOT ANNOTATED CDS; Genomic_DNA.
DR   PIR; B36733; B36733.
DR   EMBL; S63811; AAB20286.1; -; Genomic_DNA.
DR   PDB; 1DKL; X-ray; A/B=23-432.
DR   PDB; 1DKM; X-ray; A=23-432.
DR   PDB; 1DKN; X-ray; A=23-432.
DR   PDB; 1DKO; X-ray; A=23-432.
DR   PDB; 1DKP; X-ray; A=23-432.
DR   PDB; 1DKQ; X-ray; A=23-432.
DR   SWISS-2DPAGE; P07102; COLI.
DR   GenomeReviews; U00096_GR; b0980.
DR   EchoBASE; EB0047; -.
DR   EcoGene; EG10049; appA.
DR   BioCyc; EcoCyc:APPA-MONOMER; -.
DR   InterPro; IPR0005560; HisAc_phsphtse.
DR   Pfam; PF00328; Acid_phosphat_A; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW   3D-structure; Complete proteome; Direct protein sequencing; Hydrolase;
KW   Multifunctional enzyme; Periplasmic; Signal.
FT   SIGNAL        1     22
FT   CHAIN        23    432       Periplasmic appA protein.
FT                                /FTId=PRO_0000023947.
FT   ACT_SITE     39     39       Nucleophile.
FT   ACT_SITE    326    326       Proton donor.
FT   DISULFID     99    130
FT   DISULFID    155    430
FT   DISULFID    200    210
FT   DISULFID    404    413
FT   CONFLICT     51     66       MQDVTPDAWPTWPVKL -> NAGCHPRRMANLAGKT (in
FT                                Ref. 3).
FT   CONFLICT     75     76       EL -> DV (in Ref. 4).
FT   CONFLICT    112    112       D -> S (in Ref. 4).
FT   STRAND       27     38
FT   STRAND       41     42
FT   HELIX        45     53
FT   TURN         54     54
FT   STRAND       56     57
FT   TURN         63     64
FT   STRAND       66     67
FT   TURN         69     69
FT   HELIX        71     90
FT   TURN         91     92
FT   STRAND       93     94
FT   STRAND       96     98
FT   TURN        102    104
FT   STRAND      105    109
FT   STRAND      111    112
FT   HELIX       113    126
FT   HELIX       127    127
FT   TURN        128    129
FT   STRAND      134    135
FT   STRAND      138    138
FT   TURN        140    141
FT   STRAND      142    142
FT   HELIX       145    147
FT   HELIX       149    153
FT   STRAND      154    155
FT   STRAND      159    169
FT   TURN        170    172
FT   HELIX       174    179
FT   TURN        180    181
FT   HELIX       182    192
FT   TURN        193    193
FT   HELIX       194    196
FT   HELIX       197    197
FT   STRAND      198    201
FT   STRAND      202    202
FT   STRAND      204    204
FT   TURN        206    207
FT   STRAND      208    208
FT   HELIX       212    215
FT   STRAND      216    216
FT   STRAND      220    222
FT   STRAND      224    225
FT   STRAND      227    229
FT   STRAND      229    231
FT   HELIX       232    249
FT   HELIX       250    250
FT   TURN        250    253
FT   STRAND      252    253
```

```
FT   HELIX         254    257
FT   TURN          258    258
```

```
Qy   Best Local Similarity   98.1%;   Pred. No. 9.3e-162;
     Matches  424;  Conservative  0;  Mismatches  8;  Indels  0;  Gaps  0;

Qy         1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy        61 TWPVKLGELITPRGGELIAYLGHYWRQRLVADGLLPKCCGCPQSGQVAIIADVDERTRKTGE 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 TWPVKLGELITPRGGELIAYLGHYQRQRLVADGLLAKKGGCPQSGQVAIIADVDERTRKTGE 120

Qy       121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy       181 YQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 RQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy       241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy       301 LTPHPPQKQAYGVTLPTSVLPIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 LTPHPPQKQAYGVTLPTSVLPIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy       361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy       421 IVNEARIPACSL 432
             ||||||||||||
Db       421 IVNEARIPACSL 432
```

```
RESULT 2
ID   Q3Z3B1_SHISS   PRELIMINARY;   PRT;   432 AA.
AC   Q3Z3B1;
DT   27-SEP-2005, integrated into UniProtKB/TrEMBL.
DT   27-SEP-2005, sequence version 1.
DT   07-FEB-2006, entry version 5.
DE   Phosphoanhydride phosphorylase.
GN   Name=appA; OrderedLocusNames=SSO_0987; ORFNames=SSO_0987;
OS   Shigella sonnei (strain Ss046).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=300269;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   PubMed=16275786; DOI=10.1093/nar/gki954;
RA   Yang F., Yang J., Zhang X., Chen L., Jiang Y., Yan Y., Tang X.,
RA   Wang J., Xiong Z., Dong J., Xue Y., Zhu Y., Xu X., Sun L., Chen S.,
RA   Nie H., Peng J., Xu J., Wang Y., Yuan Z., Wen Y., Yao Z., Shen Y.,
RA   Qiang B., Hou Y., Yu J., Jin Q.;
RT   "Genome dynamics and diversity of Shigella species, the etiologic
RT   agents of bacillary dysentery.";
RL   Nucleic Acids Res. 33:6445-6458(2005).
```

```
EMBL; CP000038; AAZ87721.1; -; Genomic DNA.
GO; GO:0003993; F:acid phosphatase activity; IEA.
InterPro; IPR005560; HisAc_Phphtse.
Pfam; PF00328; Acid_phosphat_A; 1.
PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
```

```
KW   Complete proteome.
SQ   SEQUENCE   432 AA;   47057 MW;   6510C6C57917 7F11 CRC64;

Qy   Best Local Similarity   98.1%;   Pred. No. 9.3e-162;
     Matches  424;  Conservative  0;  Mismatches  8;  Indels  0;  Gaps  0;

Qy         1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy        61 TWPVKLGELITPRGGELIAYLGHYQRQRLVADGLLPKCCGCPQSGQVAIIADVDERTRKTGE 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 TWPVKLGELITPRGGELIAYLGHYQRQRLVADGLLAKKGGCPQSGQVAIIADVDERTRKTGE 120

Qy       121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy       181 YQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 RQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy       241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy       301 LTPHPPQKQAYGVTLPTSVLPIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 LTPHPPQKQAYGVTLPTSVLPIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy       361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy       421 IVNEARIPACSL 432
             ||||||||||||
Db       421 IVNEARIPACSL 432
```

```
RESULT 3
ID   Q8GN88_ECOLI   PRELIMINARY;   PRT;   432 AA.
AC   Q8GN88;
DT   01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT   01-MAR-2003, sequence version 1.
DT   07-FEB-2006, entry version 11.
DE   AppA.
GN   Name=appA;
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Zhang Z., He J., Yao B., Zhou Y., Chen Y., Yi Y.;
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Chen Y., Zhu Z., Zhang Z., He J.;
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
```

```
EMBL; AF537219; AAN28334.1; -; Genomic_DNA.
HSSP; P07102; 1DKM.
SMR; Q8GN88; 23-432.
GO; GO:0003993; F:acid phosphatase activity; IEA.
InterPro; IPR005560; HisAc_Phphtse.
Pfam; PF00328; Acid_phosphat_A; 1.
PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
```

```
Qy   Query Match
SQ   SEQUENCE   432 AA;   Score 2182;   DB 2;   Length 432;

Qy   Best Local Similarity   98.1%;   Pred. No. 9.3e-162;
     Matches  424;  Conservative  0;  Mismatches  8;  Indels  0;  Gaps  0;

Qy         1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy        61 TWPVKLGELITPRGGELIAYLGHYQRQRLVADGLLPKCCGCPQSGQVAIIADVDERTRKTGE 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 TWPVKLGELITPRGGELIAYLGHYQRQRLVADGLLAKKGGCPQSGQVAIIADVDERTRKTGE 120

Qy       121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy       181 YQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 RQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy       241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy       301 LTPHPPQKQAYGVTLPTSVLPIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 LTPHPPQKQAYGVTLPTSVLPIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy       361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy       421 IVNEARIPACSL 432
             ||||||||||||
Db       421 IVNEARIPACSL 432
```

SQ   SEQUENCE   432 AA; 47056 MW; 5B355D76E7377737 CRC64;

Query Match   96.5%; Score 2178; DB 2; Length 432;
Best Local Similarity   97.9%; Pred. No. 1.9e-161;
Matches 423; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

```
Qy    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60

Qy   61  TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
         ||||||||  ||||||||||||| |||||||||| || |||||||||||||||||||||
Db   61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKGCPQSGQVAIIADVDEATRKTGE  120

Qy  121  APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
         ||||||||||||||||||||||||| |||||||||||||||||||||| |||||||||||
Db  121  APAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180

Qy  181  YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240

Qy  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLLDLIKTA 300
         |||||||||||||||||||||||||||||||||||  |||||||||||||||||||||||
Db  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLLDLIKTA  300

Qy  301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGGELVPERW 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGGELVPERW  360

Qy  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420

Qy  421  IVNEARIPACSL 432
         ||||||||||||
Db  421  IVNEARIPACSL 432
```

RESULT 4
ID   Q8RKD6_ECOLI   PRELIMINARY;   PRT;   432 AA.
AC   Q8RKD6;
DT   01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT   01-JUN-2002, sequence version 1.
DT   07-FEB-2006, entry version 13.
DE   Periplasmic phosphoanhydride phosphohydrolase.
GN   Name=appA;
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=93054596; PubMed=1429631;
RA   Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M-M.,
RA   van Etten R.L.;
RT   "Overexpression, site-directed mutagenesis, and mechanism of
RT   Escherichia coli acid phosphatase.";
RL   J. Biol. Chem. 267:22830-22836(1992).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=93068616; PubMed=2168385;
RA   Dassa J., Marck C., Boquet P.L.;
RT   "The complete nucleotide sequence of the Escherichia coli gene appA
RT   reveals significant homology between pH 2.5 acid phosphatase and
RT   glucose-1-phosphatase.";
RL   J. Bacteriol. 172:5497-5500(1990).
DR   HSSP; P07102; 1DKM.
DR   SMR; Q8RKD6; 23-432.
DR   GO; GO:0003993; F:acid phosphatase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   InterPro; IPR000560; HisAc_phphtse.
DR   Pfam; PF00328; Acid_phosphat_A; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW   Hydrolase.
SQ   SEQUENCE   432 AA; 46972 MW; AF86C41EA6193AC5 CRC64;

DR   EMBL; L03374; AAA00006.1; -; Genomic_DNA.

RESULT 5
ID   Q8RKD7_ECOLI   PRELIMINARY;   PRT;   432 AA.
AC   Q8RKD7;
DT   01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT   01-JUN-2002, sequence version 1.
DT   07-FEB-2006, entry version 13.
DE   Periplasmic phosphoanhydride phosphohydrolase.
GN   Name=appA;
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=93054596; PubMed=1429631;
RA   Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M-M.,

Query Match   96.4%; Score 2176; DB 2; Length 432;
Best Local Similarity   97.9%; Pred. No. 2.7e-161;
Matches 423; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

SEQUENCE   432 AA; 46972 MW; AF86C41EA6193AC5 CRC64;

```
Qy    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60

Qy   61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKGCPQSGQVAIIADVDEATRKTGE  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKGCPQSGQVAIIADVDEATRKTGE  120

Qy  121  AFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  AFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180

Qy  181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240

Qy  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLLDLIKTA  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLLDLIKTA  300

Qy  301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGGELVPERW  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGGELVPERW  360

Qy  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420

Qy  421  IVNEARIPACSL 432
         ||||||||||||
Db  421  IVNEARIPACSL 432
```

RA   van Etten R.L.;
RT   "Overexpression, site-directed mutagenesis, and mechanism of
RT   Escherichia coli acid phosphatase.";
RL   J. Biol. Chem. 267:22830-22836(1992).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=93068616; PubMed=2168385;
RA   Dassa J., Marck C., Boquet P.L.;
RT   "The complete nucleotide sequence of the Escherichia coli gene appA
RT   reveals significant homology between pH 2.5 acid phosphatase and
RT   glucose-1-phosphatase.";
RL   J. Bacteriol. 172:5497-5500(1990).
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=93054596; PubMed=1429631;
RA   Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M-M.,

RESULT 6

```
ID   Q8RKD8_ECOLI    PRELIMINARY;    PRT;    432 AA.
AC   Q8RKD8;
DT   01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT   01-JUN-2002, sequence version 1.
DT   07-FEB-2006, entry version 13.
DE   Periplasmic phosphoanhydride phosphohydrolase.
GN   Name=appA;
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=90368616; PubMed=2168385;
```

```
RA   van Etten R.L.;
RT   "Overexpression, site-directed mutagenesis, and mechanism of
RT   Escherichia coli acid phosphatase.";
RL   J. Biol. Chem. 267:22830-22836(1992).
```

```
DR   EMBL; L03373; AAA00005.1; -; Genomic_DNA.
DR   HSSP; P07102; 1DKM.
DR   SMR; Q8RKD7; 23-432.
DR   GO; GO:0003993; F:acid phosphatase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   InterPro; IPR000560; HisAc_phsphtse.
DR   Pfam; PF00328; Acid_phosphat_A; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW   Hydrolase.
SQ   SEQUENCE   432 AA;   46972 MW;   755D5E4B1AD916A6 CRC64;

Query Match   96.4%;   Score 2176;   DB 2;   Length 432;
Best Local Similarity   97.9%;   Pred. No. 2.7e-161;
Matches 423;   Conservative 0;   Mismatches 9;   Indels 0;   Gaps 0;

Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
         |||||||.||||||||||||||.||||||||||.||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQAQRLVADGLLAKGCPQSGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy  181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
         |||||||||||||||||||||||||||||||||||||||||||||||.|||||||.|||||
Db  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy  241 EIFLLQQAQGMPEPGPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy  421 IVNEARIPACSL 432
         ||||||||||||
Db  421 IVNEARIPACSL 432
```

```
RA   Dassa J., Marck C., Boquet P.L.;
RT   "The complete nucleotide sequence of the Escherichia coli gene appA
RT   reveals significant homology between pH 2.5 acid phosphatase and
RT   glucose-1-phosphatase.";
RT   J. Bacteriol. 172:5497-5500(1990).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=90545596; PubMed=1429631;
RA   Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
RA   van Etten R.L.;
RT   "Overexpression, site-directed mutagenesis, and mechanism of
RT   Escherichia coli acid phosphatase.";
RL   J. Biol. Chem. 267:22830-22836(1992).
```

```
DR   EMBL; L03372; AAA00004.1; -; Genomic_DNA.
DR   HSSP; P07102; 1DKM.
DR   SMR; Q8RKD8; 23-432.
DR   GO; GO:0003993; F:acid phosphatase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   InterPro; IPR000560; HisAc_phsphtse.
DR   Pfam; PF00328; Acid_phosphat_A; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW   Hydrolase.
SQ   SEQUENCE   432 AA;   46972 MW;   9A85536B57PCCFB5 CRC64;

Query Match   96.4%;   Score 2176;   DB 2;   Length 432;
Best Local Similarity   97.9%;   Pred. No. 2.7e-161;
Matches 423;   Conservative 0;   Mismatches 9;   Indels 0;   Gaps 0;

Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
         |||||||.||||||||||||||.||||||||||.||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQAQRLVADGLLAKGCPQSGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy  181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
         |||||||||||||||||||||||||||||||||||||||||||||||.|||||||.|||||
Db  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy  241 EIFLLQQAQGMPEPGPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy  421 IVNEARIPACSL 432
         ||||||||||||
Db  421 IVNEARIPACSL 432
```

RESULT 7

```
ID   Q8RKE0_ECOLI    PRELIMINARY;    PRT;    432 AA.
AC   Q8RKE0;
DT   01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT   01-JUN-2002, sequence version 1.
DT   07-FEB-2006, entry version 13.
DE   Periplasmic phosphoanhydride phosphohydrolase.
```

GN   Name=appA;
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RN   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=93054596; PubMed=1429631;
RX   MEDLINE=90368616; PubMed=2168385;
RA   Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
RA   van Etten R.L.;
RT   "Overexpression, site-directed mutagenesis, and mechanism of
RA   Dassa J., Marck C., Boquet P.L.;
RT   "The complete nucleotide sequence of the Escherichia coli gene appA
RT   reveals significant homology between pH 2.5 acid phosphatase and
RT   glucose-1-phosphatase.";
RL   J. Bacteriol. 172:5497-5500(1990).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=93054596; PubMed=1429631;
RA   Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
RA   van Etten R.L.;
RT   "Overexpression, site-directed mutagenesis, and mechanism of
RT   Escherichia coli acid phosphatase.";
RL   J. Biol. Chem. 267:22830-22835(1992).
CC   ---
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---
DR   EMBL; L03370; AAA00002.1; -; Genomic_DNA.
DR   HSSP; P07102; 1DKM.
SMR;   Q8RKE0; 23-432.
GO; GO:0003993; F:acid phosphatase activity; IEA.
GO; GO:0016787; F:hydrolase activity; IEA.
InterPro; IPR000560; HiBAc_phosphtse.
Pfam; PF00328; Acid_phosphat_A; 1.
PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW   Hydrolase.
SQ   SEQUENCE   432 AA; 46972 MW; 5BBD6632D4682EABF CRC64;

Best Local Similarity   96.4%;
Matches   423;   Conservative   0;   Mismatches   9;   Indels   0;   Gaps   0;

Qy   1   MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1   MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60

Qy   61   TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61   TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy   121   AFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121   AFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy   181   YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181   RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy   241   EIFLLQQAQGMPEPGWGNRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241   EIFLLQQAQGMPEPGWGNRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy   301   LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301   LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy   361   RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361   RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy   421   IVNEARIPACSL 432
            ||||||||||||
Db   421   IVNEARIPACSL 432

RESULT 8
Q8RKD9_ECOLI
ID   Q8RKD9_ECOLI   PRELIMINARY;   PRT;   432 AA.
AC   Q8RKD9;
DT   01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT   01-JUN-2002, sequence version 1.
DT   07-FEB-2006, entry version 13.
DE   Periplasmic phosphoanhydride phosphohydrolase.
GN   Name=appA;
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=90368616; PubMed=2168385;
RA   Dassa J., Marck C., Boquet P.L.;
RT   "The complete nucleotide sequence of the Escherichia coli gene appA
RT   reveals significant homology between pH 2.5 acid phosphatase and
RT   glucose-1-phosphatase.";
RL   J. Bacteriol. 172:5497-5500(1990).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=93054596; PubMed=1429631;
RA   Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
RA   van Etten R.L.;
RT   "Overexpression, site-directed mutagenesis, and mechanism of
RT   Escherichia coli acid phosphatase.";
RL   J. Biol. Chem. 267:22830-22835(1992).
CC   ---
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   ---
DR   EMBL; L03371; AAA00003.1; -; Genomic_DNA.
DR   HSSP; P07102; 1DKQ.
SMR;   Q8RKD9; 23-432.
GO; GO:0003993; F:acid phosphatase activity; IEA.
GO; GO:0016787; F:hydrolase activity; IEA.
InterPro; IPR000560; HiBAc_phosphtse.
Pfam; PF00328; Acid_phosphat_A; 1.
PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW   Hydrolase.
SQ   SEQUENCE   432 AA; 47034 MW; 9F29B9DF9C368175 CRC64;

Best Local Similarity   96.3%;
Matches   423;   Conservative   1;   Mismatches   8;   Indels   0;   Gaps   0;

Qy   1   MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1   MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60

Qy   61   TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
            |||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db   61   TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy   121   AFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121   AFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy   181   YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181   RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy   241   EIFLLQQAQGMPEPGWGNRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241   EIFLLQQAQGMPEPGWGNRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy   301   LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301   LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy   361   RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

```
RESULT 9
Q8RKD5_ECOLI
ID   Q8RKD5_ECOLI    PRELIMINARY;     PRT;   432 AA.
AC   Q8RKD5;
DT   01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT   01-JUN-2002, sequence version 1.
DT   07-FEB-2006, entry version 13.
DE   Periplasmic phosphoanhydride phosphohydrolase.
GN   Name=appA;
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=90368616; PubMed=2168385;
RA   Dassa J., Marck C., Boquet P.L.;
RT   "The complete nucleotide sequence of the Escherichia coli gene appA
RT   reveals significant homology between pH 2.5 acid phosphatase and
RT   glucose-1-phosphatase.";
RL   J. Bacteriol. 172:5497-5500(1990).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=93054556; PubMed=1429631;
RA   Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
RA   van Etten R.L.;
RT   "Overexpression, site-directed mutagenesis, and mechanism of
RT   Escherichia coli acid phosphatase.";
RL   J. Biol. Chem. 267:22830-22836(1992).
```

```
DR   EMBL; L03375; AAA00007.1; -; Genomic_DNA.
DR   HSSP; P07102; 1DXM.
DR   SMR; Q8RKD5; 23-432.
DR   GO; GO:0003993; F:acid phosphatase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   InterPro; IPR005560; HiAc_phsphtse.
DR   Pfam; PF00328; Acid_phosphat_A; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
KW   Hydrolase.
SQ   SEQUENCE   432 AA;  46991 MW;  951F393EA9A1A47C CRC64;

Query Match             96.2%;  Score 2172;  DB 2;  Length 432;
Best Local Similarity   97.9%;
Matches  423;  Conservative   0;  Mismatches  9;  Indels  0;  Gaps  0;

Qy          1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy         61 TWPVKLGELIPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
              |||||| ||||||||||||||    ||||||||  ||||||||||||||||||||||||||
Db         61 TWPVKLGWLIPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy        121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db        121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy        181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy        241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300

Qy        301 LTPHPPQKQAYGVTLPTSVLPIAGADTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LTPHPPQKQAYGVTLPTSVLPIAGADTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy        361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy        421 IVNEARIPACSL 432
              ||||||||||||
Db        421 IVNEARIPACSL 432
```

```
RESULT 10
Q31YP3_SHIBS
ID   Q31YP3_SHIBS    PRELIMINARY;     PRT;   432 AA.
AC   Q31YP3;
DT   06-DEC-2005, integrated into UniProtKB/TrEMBL.
DT   06-DEC-2005, sequence version 1.
DT   07-FEB-2006, entry version 4.
DE   Phosphoanhydride phosphorylase.
GN   Name=appA; OrderedLocusNames=SBO_2250;
OS   Shigella boydii serotype 4 (strain Sb227).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=300268;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   PubMed=16275786; DOI=10.1093/nar/gki954;
RA   Yang F., Yang J., Zhang X., Chen L., Jiang Y., Yan Y., Tang X.,
RA   Wang J., Xiong Z., Dong J., Xue Y., Zhu Y., Xu X., Sun L., Chen S.,
RA   Nie H., Peng J., Xu J., Wang Y., Yuan Z., Wen Y., Yao Z., Shen Y.,
RA   Qiang B., Hou Y., Yu J., Jin Q.;
RT   "Genome dynamics and diversity of Shigella species, the etiologic
RT   agents of bacillary dysentery.";
RL   Nucleic Acids Res. 33:6445-6458(2005).
```

```
DR   EMBL; CP000036; ABB66815.1; -; Genomic_DNA.
DR   GO; GO:0003993; F:acid phosphatase activity; IEA.
KW   Complete proteome.
SQ   SEQUENCE   432 AA;  47064 MW;  8569022BA28D9C73 CRC64;

Query Match             95.4%;  Score 2154;  DB 2;  Length 432;
Best Local Similarity   97.0%;
Matches  419;  Conservative   0;  Mismatches  13;  Indels  0;  Gaps  0;

Qy          1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy         61 TWPVKLGELIPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
              |||||| |||||||||||| |||    ||||||||  |||||||||||||||||||||||
Db         61 TWPVKLGWLIPRGGELIAYFGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy        121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db        121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy        181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy        241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIMAA 300

Qy        301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNIANLGGALELNWTLPGQPDNTPPGGELVFERW 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNIANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy        361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy        421 IVNEARIPACSL 432
              ||||||||||||
Db        421 IVNEARIPACSL 432
```

RESULT 11
Q83RW2_SHIFL

```
ID   Q83RW2_SHIFL            PRELIMINARY;      PRT;   432 AA.
AC   Q83RW2;
DT   01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT   05-JUL-2005, sequence version 15.
DT   07-FEB-2006, entry version 15.
DE   Phosphoanhydride phosphorylase; pH 2.5 acid phosphatase.
GN   Name=appA; OrderedLocusNames=SF0982;
OS   Shigella flexneri.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=623;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=301 / Serotype 2a;
RX   MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA   Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA   Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA   Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA   Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA   Yu J.;
RT   "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT   through comparison with genomes of Escherichia coli K12 and O157.";
RL   Nucleic Acids Res. 30:4432-4441(2002).
CC   -----------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   -----------------------------------------------------------------------
DR   EMBL; AE005674; AAN42610.2; -; Genomic_DNA.
DR   SMR; Q83RW2; 23-432.
DR   BioCyc; SFLE198214:AAN42610.1-MONOMER; -.
DR   GO; GO:0003993; F:acid phosphatase activity; IEA.
DR   InterPro; IPR000560; HisAc_phsphtse.
DR   Pfam; PF00328; Acid_phosphat_A; 2.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW   Complete proteome.
SQ   SEQUENCE   432 AA;  47106 MW;  25093A53392B9C18C CRC64;

Query Match             95.2%;  Score 2149;  DB 2;  Length 432;
Best Local Similarity   96.8%;  Pred. No. 3.5e-159;
Matches  418;  Conservative   1;  Mismatches  13;  Indels  0;  Gaps  0;

Qy      1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy     61 TWPVKLGWLTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120

Qy    121 AFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180

Qy    181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
          ||| ||||||||||||||||| ||||||||||||||||||||||| ||||||||||||||
Db    181 RQTVFRELERVLNFPQSNLCLNREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy    241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIMA 300

Qy    301 LITPHPPQKQAYGVTLPTSVLFIAGHDTNLANIGGALELNWTLPGQPDNTPGGELVFERW 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 LITPHPPQKQAYGVTLPTSVLFIAGHDTNLANIGGALELNWTLPGQPDNTPGGELVFERW 360

Qy    361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy    421 IVNEARIPACSL 432
          ||||||||||||
Db    421 IVNEARIPACSL 432
```

RESULT 12
Q6RK08_ECOLI

```
ID   Q6RK08_ECOLI            PRELIMINARY;      PRT;   432 AA.
AC   Q6RK08;
DT   05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT   05-JUL-2004, sequence version 1.
DT   07-FEB-2006, entry version 8.
DE   Acid phosphatase/phytase 2.
GN   Name=appA2;
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=99194564; PubMed=10092520; DOI=10.1006/bbrc.1999.0361;
RA   Rodríguez E., Han Y., Lei X.G.;
RT   "Cloning, sequencing, and expression of an Escherichia coli acid
RT   phosphatase/phytase gene (appA2) isolated from pig colon.";
RL   Biochem. Biophys. Res. Commun. 257:117-123(1999).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Lei X.G.;
RL   Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.
CC   -----------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   -----------------------------------------------------------------------
DR   EMBL; AY496073; AAR87658.1; -; Genomic_DNA.
DR   SMR; Q6RK08; 23-432.
DR   GO; GO:0003993; F:acid phosphatase activity; IEA.
DR   InterPro; IPR000560; HisAc_phsphtse.
DR   Pfam; PF00328; Acid_phosphat_A; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
FT   CHAIN        20    432       acid phosphatase/phytase 2.
SQ   SEQUENCE   432 AA;  47042 MW;  71B05EED2EA2674 CRC64;

Query Match             95.1%;  Score 2147;  DB 2;  Length 432;
Best Local Similarity   96.8%;  Pred. No. 5.1e-159;
Matches  418;  Conservative   0;  Mismatches  14;  Indels  0;  Gaps  0;

Qy      1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy     61 TWPVKLGWLTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
          ||||||| |||||||||||||||||||||||||| |||||||||||||||||||||||||
Db     61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy    121 AFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    121 AFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy    181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
          ||||||||||||||||||||| ||||||||||||||||||||||| ||||||||||||||
Db    181 RQTAFRELERVLNFPQSNLCLNREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy    241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIMA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIMA 300

Qy    301 LITPHPPQKQAYGVTLPTSVLFIAGHDTNLANIGGALELNWTLPGQPDNTPGGELVFERW 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 LITPHPPQKQAYGVTLPTSVLFIAGHDTNLANIGGALELNWTLPGQPDNTPGGELVFERW 360

Qy    361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy    421 IVNEARIPACSL 432
          ||||||||||||
Db    421 IVNEARIPACSL 432
```

RESULT 13

```
Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDJLLQRTPEVARSRATPLLLDLIKTA 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDJLLQRTPEVARSRATPLLLDLIMAA 300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLLPGQPDNTPPGGELVFERW 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLLPGQPDNTPPGGELVFERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy  421 IVNEARIPACSL 432
        ||||||||||||
Db  421 IVNEARIPACSL 432
```

```
ID   Q7UD08_SHIFL   PRELIMINARY;   PRT;   432 AA.
AC   Q7UD08;
DT   01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT   01-OCT-2003, sequence version 1.
DT   07-MAR-2006, entry version 12.
DE   Phosphoanhydride phosphorylase.
GN   Name=appA; OrderedLocusNames=S1048; ORFNames=S_1048;
OS   Shigella flexneri.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=623;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX   MEDLINE=22590274; PubMed=12704152;
RA   DOI=10.1128/IAI.71.5.2775-2786.2003;
RA   Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA   Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA   Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA   Schwartz D.C., Blattner F.R.;
RT   "Complete genome sequence and comparative genomics of Shigella
RT   flexneri serotype 2a strain 2457T.";
RL   Infect. Immun. 71:2775-2786(2003).
-----------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
-----------------------------------------------------------------------
DR   EMBL; AE014073; AAP16495.1; -; Genomic_DNA.
DR   HSSP; P07102; 1DKL.
DR   SMR; Q7UD08; 23-432.
DR   GenomeReviews; AE014073_GR; S1048.
DR   GO; GO:0003993; F:acid phosphatase activity; IEA.
DR   InterPro; IPR000560; HisAc_phphtse.
DR   Pfam; PF00328; Acid_phosphat_A; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN_1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
SQ   SEQUENCE   432 AA;  47136 MW;  200422398C66DFC CRC64;
```

Query Match   95.0%; Score 2145; DB 2; Length 432;
Best Local Similarity   96.5%;
Matches 417; Conservative 14; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 MKAILIPPLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP 60
        |||||||||||||||||||| : |||||||||:||||||||||||||||||||||||||
Db    1 MKAILIPPLSLLIPLTPQSAFAQSEPELKLRSMVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTKTGE 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKGCPQSAQVAIIADVDERTKTGE 120

Qy  121 APAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 APTAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILCRAGGSIADFTGH 180
```

RESULT 14

```
Qy  181 YQTAFRELERVLNFPQSNLCLKREKQDSSCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
        |||||||||||||||||||: ||||:::|| ::||:| |||||||| :|||:|||||||
Db  181 RQTVFRELERVLNFPQSNLCLNREKQDESSCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDJLLQRTPEVARSRATPLLLDLIKTA 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLLDLIMAA 300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLLPGQPDNTPPGGELVFERW 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLLPGQPDNTPPGGELVFERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy  421 IVNEARIPACSL 432
        ||||||||||||
Db  421 IVNEARIPACSL 432
```

```
ID   Q7AFW3_ECO57   PRELIMINARY;   PRT;   434 AA.
AC   Q7AFW3;
DT   05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT   05-JUL-2004, sequence version 1.
DT   07-FEB-2006, entry version 10.
DE   Phosphoanhydride phosphorylase.
GN   OrderedLocusNames=ECs1136;
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=83334;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX   MEDLINE=21562331; PubMed=11258796; DOI=10.1093/dnares/8.1.11;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
-----------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
-----------------------------------------------------------------------
DR   EMBL; BA000007; BAB34559.1; -; Genomic_DNA.
DR   SMR; Q7AFW3; 25-434.
DR   GO; GO:0003993; F:acid phosphatase activity; IEA.
DR   InterPro; IPR000560; HisAc_phphtse.
DR   Pfam; PF00328; Acid_phosphat_A; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
SQ   SEQUENCE   434 AA;  47337 MW;  F197D71869F9C4 CRC64;
```

Query Match   95.0%; Score 2144; DB 2; Length 434;
Best Local Similarity   96.5%;
Matches 419; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

```
Qy    1 MKAILIPPLSLLIPLTPQSAFAQS--EPELKLESVVISRHGVRAPTKATQLMQDVTPDA 58
        |||||||||||||||||||||||:  :|||||||||||||||||||||||||||||||:
Db    1 MKAILIPPLSLLIPLTPQSAFAQSEPEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDA 60

Qy   59 WPTWPVKLGWLTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTKT 118
        |||||||||||||||||||||||||:|||||||||||:|||:||||||||||||||||
Db   61 WPNWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLTKKGCPQPGQVAIIADVDERTKT 120

Qy  119 GEAPAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFT 178
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GEAPAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFT 178
```

```
Db  121  GEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVGQLDNANVTDAILSRAGGSTADFT  180
Qy  179  GHYQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASM  238
         ||| ||||||||||||||||||| ||||||||||||||||||||| |||||||||||||
Db  181  GHRQTAFRELERVLNFPQSNLCLNREKQDESCSLTQALPSELKVSADNVSLTGAVSLASM  240
Qy  239  LTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIK  298
         |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db  241  LTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIM  300
Qy  299  TALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPCQPDNTPPGGELVFE  358
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  IALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPCQPDNTPPGGELVFE  360
Qy  359  RWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGF  418
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  RWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGF  420
Qy  419  TQIVNEARIPACSL  432
         ||||||||||||||
Db  421  TQIVNEARIPACSL  434
```

RESULT 15
Q8XC29 ECO57

```
ID   Q8XC29;        PRELIMINARY;      PRT;      444 AA.
AC   Q8XC29;
DT   01-MAR-2002, integrated into UniProtKB/TrEMBL.
DT   05-JUL-2004, sequence version 2.
DT   07-FEB-2006, entry version 17.
DE   Phosphoanhydride phosphorylase; pH 2.5 acid phosphatase; periplasmic.
GN   Name=appA; OrderedLocusNames=21397;
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=83334;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX   MEDLINE=2107495; PubMed=11206551; DOI=10.1038/35054089;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
CC   -----------------------------------------------------------------------
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   -----------------------------------------------------------------------
DR   EMBL; AE005174; AAG55528.1; -; Genomic_DNA.
DR   PIR; D85633; D85633.
DR   PIR; H90770; H90770.
DR   HSSP; P07102; 1DKL.
     SMR; Q8XC29; 35-444.
DR   BioCyc; ECOL83334-1:ECS1136-MONOMER; -.
DR   GO; GO:0003993; F:acid phosphatase activity; IEA.
DR   InterPro; IPR000560; H8Ac_phsphtse.
DR   Pfam; PF00328; Acid_phosphat_A; 1.
DR   PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR   PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW   Complete proteome.
SQ   SEQUENCE   444 AA;  48388 MW;  A9AA7B6653AF247B CRC64;
```

```
Query Match            95.0%;  Score 2144;  DB 2;  Length 444;
Best Local Similarity  96.5%;  Pred. No. 9e-159;
Matches  419;  Conservative   0;  Mismatches 13;  Indels 2;  Gaps 1;
```

```
Qy    1  MKAILIPFLSLLIPLPTPQSAFAQS--EPELKLESVVISRGVRAPTKATQLMQDVTPDA  58
         ||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db   11  MKAILIPFLSLLIPLPTPQSAFAQSEPEPELKLESVVISRGVRAPTKATQLMQDVTPDA  70
```

```
Qy   59  WPTWPVKLGBLTPRGGELIIAYLGHYWRQRLVADGLLPKCCGCPQSGQVAIIADVDERTRKT  118
         ||| ||||| ||||||||| ||||| ||||||||||    | ||| ||||||||||||||||
Db   71  WPNWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLTKKGCQPQGQVAIIADVDERTRKT  130
Qy  119  GEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVGQLDNANVTDAILERAGGSTADFT  178
         ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db  131  GEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVGQLDNANVTDAILSRAGGSTADFT  190
Qy  179  GHYQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASM  238
         ||| |||||||||||||||||||| ||||||||||||||||||||||| |||||||||||
Db  191  GHRQTAFRELERVLNFPQSNLCLNREKQDESCSLTQALPSELKVSADNVSLTGAVSLASM  250
Qy  239  LTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIK  298
         |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db  251  LTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIM  310
Qy  299  TALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPCQPDNTPPGGELVFE  358
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  311  IALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPCQPDNTPPGGELVFE  370
Qy  359  RWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGF  418
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  371  RWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGF  430
Qy  419  TQIVNEARIPACSL  432
         ||||||||||||||
Db  431  TQIVNEARIPACSL  444
```

Search completed: June 13, 2006, 10:28:33
Job time : 190.248 secs

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

Run on: June 14, 2006, 12:41:52 ; Search time 220.459 Seconds
(without alignments)
5499.798 Million cell updates/sec

Title: US-10-601-319-10
Perfect score: 2258
Sequence: 1 MKAILIPFLSLLIPLTPQSA.........CSLAGFTQIVNEARIPACSL 432

Scoring table: BLOSUM62
Xgapop 10.0 ; Xgapext 0.5
Ygapop 10.0 ; Ygapext 0.5
Fgapop 6.0 ; Fgapext 7.0
Delop 6.0 ; Delext 7.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Command line parameters:
1: -MODEL=frame+_p2n.model -DEV=xlp
2: /EMC_Celerra_SIDS3/ptodata/2/ina/1_COMB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/ina/5_COMB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/ina/6A_COMB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/ina/6B_COMB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/ina/7_COMB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/ina/H_COMB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP_COMB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/ina/PCTUS_COMB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/ina/RE_COMB.seq:*
-Q=/abss/ABSSWEB_spool/US10601319/runat_12062006_180055_27578/app_query.fasta_1
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cd1
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-HOST=abss03p -USER=US10601319 @CGN_1_1_381 @runat_12062006_180055_27578
-NCPU=6 -ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2182	96.6	1901	3	US-09-866-379-7	Sequence 7, Appli
2	2182	96.6	1901	3	US-09-866-379-9	Sequence 9, Appli
3	2181	96.6	1901	3	US-09-866-379-5	Sequence 5, Appli
4	2175	96.3	1901	3	US-09-866-379-6	Sequence 6, Appli
5	2171	96.1	1323	3	US-09-259-214-1	Sequence 1, Appli
6	2171	96.1	1323	3	US-09-318-528-1	Sequence 1, Appli
7	2171	96.1	1323	3	US-09-291-931-1	Sequence 1, Appli
8	2171	96.1	1323	3	US-09-580-515-1	Sequence 1, Appli
9	2171	96.1	1323	3	US-09-866-379-1	Sequence 1, Appli
10	2164	95.8	1486	3	US-09-715-477-2	Sequence 2, Appli
11	2147	95.1	1489	3	US-09-540-149A-9	Sequence 9, Appli
12	2147	95.1	1489	4	US-10-265-041A-9	Sequence 9, Appli
13	2144	95.1	1486	3	US-10-715-477-4	Sequence 4, Appli
14	2063.5	91.4	1272	3	US-08-910-798-1	Sequence 341, App
15	543	24.0	1266	3	US-09-489-0399A-341	Sequence 341, Ap
16	381	16.9	1569	3	US-09-489-039A-6330	Sequence 6330, Ap
17	356	15.8	5975	3	US-08-920-812-23	Sequence 23, Appl
18	356	15.8	5975	3	US-08-920-827-23	Sequence 23, Appl
19	356	15.8	5975	3	US-08-362-577C-23	Sequence 23, Appl
20	356	15.8	5975	2	US-08-920-177-23	Sequence 23, Appl
21	356	15.8	5975	3	US-08-868-435-32	Sequence 32, Appl
22	356	15.8	5975	2	US-08-920-828-23	Sequence 23, Appl
23	356	15.6	1571	3	US-08-744-231-32	Sequence 32, Appl
24	126	5.6	1571	3	US-09-044-718-10	Sequence 10, Appl
25	126	5.6	1571	3	US-09-635-504-32	Sequence 32, Appl
26	126	5.6	1571	3	US-10-062-848-10	Sequence 10, App
27	124	5.5	1404	3	US-09-684-855-166	Sequence 166, App
28	124	5.5	1404	3	US-09-488-265B-32	Sequence 32, Appl
29	117	5.2	7404	3	US-09-902-540-3115	Sequence 3115, Ap
30	117	5.2	18809	3	US-08-999-359-27	Sequence 27, Appl
31	116	5.1	1642	3	US-09-482-558A-27	Sequence 27, Appl
32	116	5.1	1642	3	US-09-902-540-1141	Sequence 1141, Ap
33	115	5.1	12508	3	US-09-655-270A-1	Sequence 1, Appl
34	115	5.1	12523	3	US-09-651-941-1	Sequence 1, Appl
35	115	5.1	12523	3	US-09-955-597-1	Sequence 1, Appl
36	114	5.0	28958	3	US-08-258-261B-6	Sequence 6, Appl
37	114	5.0	28958	2	US-08-456-837-6	Sequence 6, Appl
38	114	5.0	28958	2	US-08-457-342-6	Sequence 6, Appl
39	114	5.0	28958	2	US-08-457-646A-6	Sequence 6, Appl
40	114	5.0	28958	2	US-08-458-076A-6	Sequence 6, Appl
41	114	5.0	28958	2	US-08-764-233A-4	Sequence 4, Appl
42	114	5.0	28958	2	US-08-457-335A-6	Sequence 6, Appl
43	114	5.0	28958	3	US-08-729-214-6	Sequence 6, Appl
44	114	5.0	28958	3	US-09-028-934-6	Sequence 6, Appl
45	114	5.0	49377	2	US-08-764-233A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-866-379-7
; Sequence 7, Application US/09866379
; Patent No. 6853365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 7
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli

```
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (1)..(1901)
;  OTHER INFORMATION: n is any nucleotide
US-09-866-379-7

Alignment Scores:
Pred. No.:             5.28e-223
Score:                 2182.00
Percent Similarity:    98.1%        Length:       1901
Best Local Similarity: 98.1%        Matches:       424
Query Match:           96.6%        Conservative:    0
DB:                    3            Mismatches:      8
                                    Indels:          0
                                    Gaps:            0

US-10-601-319-10 (1-432) x US-09-866-379-7 (1-1901)

Qy      1 MetLysAlaIleLeuIleLeuProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db    188 ATGAAAGCGATCTTAATCCCATTTTATCTCTTCTGATTCCGTTAACCCCAATCTGCA 247

Qy     21 PheAlaGlnSerGluProGluLeuLeuGluSerValValIleValSerArgHisGly 40
Db    248 TTCGCTCAGAGTGAGCCGGAACTGAAAGTCGGAATTGTCAGTTCGTTCATGGT 307

Qy     41 ValArgAlaProThrArgGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db    308 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCGACGGCATGGCCA 367

Qy     61 ThrTrpProValLysGlnLeuGlyGluThrProArgGlyGlyGluIleIleAlaTyrLeu 80
Db    368 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCCGGCNGGTGGTGAGCTAATCGCCTATCTC 427

Qy     81 GlyHisTyrThrArgGlnAsnArgLeuValAlaAspGlyLeuLeuProLysCysCysPro 100
Db    428 GGACAATTACCAACGCCAGGCGTCGGTAGCCCACGGATTGCTGCCGAAAAAGGGCTGCCCG 487

Qy    101 GlnSerGlyGlnValAlaIleLeuAlaAspGlyLeuGluArgThrArgLysThrGlyGlu 120
Db    488 CAGTCTGGTCAGGTCGCGATTCTGGCAGATGGTCTGGAACGGACTCGTAAAACAGGCGAA 547

Qy    121 AlaPheAlaLeuAlaGlyLeuAlaProAspPheCysSerAlaIleThrValHisThrGlnAlaAspThr 140
Db    548 GCCTTCGCCCTGGCGGGTCTGGCACCTGACTGTGCCAATACCGTACATACCCAGGCAGATACG 607

Qy    141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAspAla 160
Db    608 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACGGGCGTTTGCCAACTGGATAACGCG 667

Qy    161 AsnValThrAspAlaIleLeuGluArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db    668 AACGTGACTGACGCGATCCTCAGCAGGGCCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 727

Qy    181 TyrGlnThrArgAlaPheArgGluLeuuArgGlnValLeuAsnPheProGlnSerAsnLeuCys 200
Db    728 CGGCAAACGGCGTTTCGGGAACTGGACGTGCTTAATTTCCGAATCAAACTTGTGC 787

Qy    201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db    788 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCACGGCAATACCATCGGAACTC 847

Qy    221 LysValSerAlaAspCysValSerLeuThrGlyAlaAlaValSerLeuValAlaSerMetLeuThr 240
Db    848 AAGGTGAGCGCCGACAAAGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTAACG 907

Qy    241 GluIlePheLeuLeuGlnIleAlaGlnGlyTyrMetProGluProGlyTyrTrpGlyTyrArgIleThr 260
Db    908 GAGATATTTCTCCTGCAACAAGCACAGGGATACATGCCCGGGCGCCGGGGTGGGGGATCACC 967

Qy    261 AspSerHisGlnThrAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
Db    968 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGGCAATTTATTGCTACACA 1027

Qy    281 ArgThrThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
```

```
Db   1028 CGCACGACGCCAGAGGTTGCCCGCAGCGCGGCCACCCCGTTATTGGATTTGATTAAGACAGCG 1087

Qy    301 LeuThrProHisProProGlnIleLysGlnGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db   1088 TTGACGCCCCATCCACCGCAAAAACAGCAGGCGTATGGGTGACATTACCCACTTCAGTGCTG 1147

Qy    321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db   1148 TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG 1207

Qy    341 ThrLeuProGlyGlnProAspAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db   1208 ACGCTTCCCGGTCAGCCCGACGATAACACCCCGCCGGGTGGTGAACTGGTGTTTGAACGCTGG 1267

Qy    361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db   1268 CGTCGGCTGAGCGATAACAGCCAGTGGATTCAGGTTTCGCTCGGTCTTCCAGACTTTACAG 1327

Qy    381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyGluValValLysLeuThr 400
Db   1328 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCGGAGGTGAAAACTGACC 1387

Qy    401 LeuAlaGlyCysGluGluGluArgAsnAlaGlnGlyTyrMetCysSerLeuValAlaGlyPheThrGln 420
Db   1388 CTGGCAGGATGTGAAGAGGCGAAATGCCAGGGCCATGGTCGTTCGTTGGTGGCAGGTTTTACCCAA 1447

Qy    421 IleValAlaAsnGluAlaArgIleProAlaCysSerLeu 432
Db   1448 ATCGTGAATGAAGCACCGCGATACCCGGCCTGCAGTTTG 1483
```

```
RESULT 2
US-09-866-379-9
; Sequence 9, Application US/09866379
; Patent No. 6855365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 9
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-9

Alignment Scores:
Pred. No.:             5.28e-223
Score:                 2182.00      Length:       1901
Percent Similarity:    98.1%        Matches:       424
                                    Conservative:    0
```

```
Best Local Similarity:  98.1%        Mismatches:  8
Query Match:            96.6%        Indels:      0
                        3           Gaps:        0

US-10-601-319-10 (1-432) x US-09-866-379-9 (1-1901)

Qy    1  MetLysAlaIleIleLeuIlePheProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla    20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  188  ATGAAAGCGGATCTTAAATCCCATTTTTACTCTCTTCTGATTCCGTTAACCCCGCAATCTGCA    247

Qy   21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly         40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  248  TTCGCTCAGAGTGAGCCCGGAGCTGAAGCTCGGAAAGTGGTGATAGTCTCACGGT            307

Qy   41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAlaTrpPro           60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  308  GTGCGTCGTCCAACCAAGGCCACGCAACTGATGCAAGATGTCACCCCAGCATGGCCA           367

Qy   61  ThrTrpProValLysGluLysGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu     80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  368  ACCTGGCCGGTAAAAGAGAAAGGGGTTGGCTGACACCCCGNGGTGGTGAGCTAATCGCCTATCTC    427

Qy   81  GlyHisTyrTrpAspGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro       100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  428  GGACATTACCAACGCCAGCGTCTGGTGGCCGACGGATTGCTGCCGAAAAAGGCTGCCCG        487

Qy  101  GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu      120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  488  CAGTCTGGTCAGGTCGCGATTATTGCCGATGTCGACGAGCGTACCCGTAAAAACAGGCGAA     547

Qy  121  AlaPheAlaAlaGlyLeuAlaProAspAspAspCysAlaIleThrValHisThrGlnAlaAspThr  140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  548  GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG       607

Qy  141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla      160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  608  TCCAGTCCCGATCCGTTATTTAAATCCTCTAAAAACTGGCGTTGCCAACTGGATACCGCG      667

Qy  161  AsnValThrAspProAlaIleIleLeuAlaAsnGlyAlaGlySerIleLeuAlaAspPhePheThrGlyHis  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  668  AACGTGACTGACCCGGCGATCCTCAGCAGGGCAGGGTCAATTGCTGACTTTACCGGGCAT      727

Qy  181  TyrGlnThrArgLeuPheArgGlyLeuGluLeuValAsnProGlnSerAsnLeuCys         200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  728  CGGCAAAACGGGCGTTTCCGGAACTGGAGGTTGTCTTAATTTCCGCAATCAAACTTGTGC     787

Qy  201  LeuLysArgGlyLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu      220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  788  CTTAAAACGTGAGAAACAGGACCAGAAGCTGTTCATTAACGCACAGGCATTACCATCGGAACTC   847

Qy  221  LysValSerAlaAspSerCysValSerSerLeuThrGlyValAlaValSerArgMetLeuThr    240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  848  AAGGTAAGCGCCGCAAAGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAAGGCTGACG       907

Qy  241  GluIlePheLeuLeuGlnGlnAlaGlnLeuGlyMetProGlyProGlyTrpGlyArgIleThr   260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  908  GAGATATTTCCTCGCAACAAGCACAGGGAAGCGCCGGGGTGGGAAGGGATCACC            967

Qy  261  AspSerHisGlnThrTrpAlaAsnThrLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln   280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  968  GATTCACACCAGTGGGAACACCTTGCTAAGTTGCATAACGCGCAATTTTTGTTTGCTACAA    1027

Qy  281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla     300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1028  CGCACGCCCAGAGGTTGCCCGCAGCCGGGCGCCACCCGTTATTAGATTTGATCAAGACAGCG   1087

Qy  301  LeuThrProHisProProGlyLysGlnLeuTyrGlyValThrThrLeuProThrSerValLeu   320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1088  TTGACGCCCCATCCACCGCAAAAACAGGGGTGACATTACCCACTTCAGTGGTG            1147

Qy  321  PheIleAlaGlyHisAspThrAsnLeuAsnLeuAsnLeuGlyGlyAlaLeuGluLeuAsnTrp   340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1148  TTTATCGCCGGACACGATACTAAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG    1207

Qy  341  ThrLeuProGlyGlnProAlaGlnAspAsnThrTyrIleGluLeuValSerSerLeuValPheGlnGlnThrLeuGln  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1208  ACGGTTCCGGGTCAGCCCGCACAGGATAACACACGCCGCGGATAACAGCCAGGTTGGCGGTGAAACTGGTTGCTGG    1267

Qy  361  ArgArgLeuSerArgAspAspAsnSerGlnTrpProGluGlnValSerArgGluValPheGlnThrGluLeuGln    380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1268  CGTCGGCGTAAGCGCGATAACGCCCAGTGGCCGGATAACAGCCAGGTTTCGGTGGCTGGTCTTCCAGGACTTTACAG   1327

Qy  381  GlnMetArgAspSerLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuLysThr   400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1328  CAGATGCGTGATAGAAAAACGCCGCTGTCATTAAATACACCGCCCGGAGAGGGTGAAACTGACC   1387

Qy  401  LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln     420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1388  CTGGCAGGATGTGAAGAGCGGAATGCAGGGCCATGTGTTCGTTGGTTTTACGCAA          1447

Qy  421  IleValAsnGluAlaArgIleProAlaCysSerLeu     432
         ||||||||||||||||||||||||||||||||||||
Db 1448  ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTG     1483
```

RESULT 3
US-09-866-379-5

; Sequence 5, Application US/09866379
; Patent No. 6853365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-5

Alignment Scores:
Pred. No.:
Score: 6.75e-223
Percent Similarity: 98.1%   Length: 1901
Best Local Similarity: 98.1%   Matches: 424
Query Match: 96.6%   Conservative: 0
                        3   Mismatches: 8
                                     Indels: 0
                                     Gaps: 0

US-10-601-319-10 (1-432) x US-09-866-379-5 (1-1901)

```
Qy    1  MetLysAlaIleIleLeuIlePheProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla    20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  188  ATGAAAGCGGATCTTAAATCCCATTTTTACTCTCTTCTGATTCCGTTAACCCCGCAATCTGCA    247
```

```
Qy    21  PheAlaGlnSerGluProGluLeuLeuLysGluLeuSerValValIleValSerArgHisGly    40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248  TTCGCTCAGAGTGAGCCGGAGCTGGAAGTGTGGTGATTGTCAGTCGTCATGGT   307

Qy    41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308  GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA   367

Qy    61  ThrTrpProValLysLeuGlyLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu    80
          |||||||||||||||||||||||||||||||||||||||||||||
Db   368  ACCTGGCCGGTAAAACTGGGTTTGACTCACCGCGGTGGTGAGCTAATCGCCTATCTC   427

Qy    81  GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro   100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428  GGACATTACCAACGCCAGCGGCTCGTAGCCGACGGACTGCTGCCGAAAAAGGGCTGCCCG   487

Qy   101  GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu   120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488  CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACACGGGAA   547

Qy   121  AlaAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisHisGlnAlaAspThr   140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548  GCCTTCGCCGCCGGCCTGGCACCTGACTGTGCCATCACCGTACATCACCAGGCAGATACG   607

Qy   141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla   160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   608  TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG   667

Qy   161  AsnValThrAspAlaIleLeuGluArgAlaGlyGlySerIleAlaAspPhePheThrGlyHis   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668  AACGTGACTGACGCGATCCTCAGCGAGCGCAGAGGGTCAATGTGCTACTTTACCGGGCAT   727

Qy   181  TyrGlnThrAlaPheAlaArgGluLeuGlyLeuValLeuAsnProMetGlnSerAsnLeuCys   200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   728  CGGCAAACGGCGTTTGCGAGACTGGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC   787

Qy   201  LeuLysArgGlyLysGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu   220
          ||||||||||||||||||||||||||||||||||||||||||||
Db   788  CTTAAACGTGGAAAACAGGAAAGCTGTTCAATACCAGGCCATTACCACCTGGAACTC   847

Qy   221  LysValSerAlaAspCysValSerLeuThrGlyAlaAlaValSerLeuMetLeuThr   240
          ||||||||||||||||||||||||||||||||||||||||||||
Db   848  AAGGTGAGCGCCGACAAATGTCTCATTAACCCGTGCGGTAAGCCTCGCATCAATGCTACG   907

Qy   241  GluIlePheLeuLeuGluGlnIleAlaGlyIleGlyMetProGlyProGlyTyrProGlyIleIleThr   260
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   908  GAGATATTTCTCCTGGAACAAGCACAGGGAATGCCCGGACCCGGGTGGGGAAGGATCACC   967

Qy   261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln   280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   968  GATTCACACCAGTGGAACACCTTGCTAAGTGTTGCATAACGCCCAATTTATTGCTACAA   1027

Qy   281  ArgThrThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1028  CGCACCACCAGAGGTTGCCCGCAGCCGCCACCCCGTTATTAGATTTGATCAAGACAGCG   1087

Qy   301  LeuThrProHisProProGlnLysGlyIleAlaTyrGlyValThrLeuProThrSerValLeu   320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1088  TTGACGCCCCATCCACCGCAAAAAACAGGCATAGCGTATGGCGTGACATTACCGACTTCTG   1147

Qy   321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp   340
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1148  TTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG   1207

Qy   341  ThrLeuProGlyProAspAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp   360
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1208  ACGCTTCCCGGTCCGGATGATAACACGCCGCCGGGTGGTGAACTGGTGTTTGAACGCTGG   1267

Qy   361  ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln   380
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1268  CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTCGTCTTCCAGACTTTACAG   1327

Qy   381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuLeuThr   400
          |||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 4
US-09-866-379-6
; Sequence 6, Application US/09866379
; Patent No. 6855365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 6
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-6

US-10-601-319-10 (1-432) x US-09-866-379-6 (1-1901)

; Alignment Scores:
Pred. No.:
Score: 2.96e-222   Length: 1901
Percent Similarity: 2175.00   Matches: 423
Best Local Similarity: 97.9%   Conservative: 0
Query Match: 97.9%   Mismatches: 9
DB: 96.3%   Indels: 0
                                     Gaps: 0

```
Qy     1  MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla    20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  ATGAAAGCGGATCTTAATCCCATTTTTATCTCTTCTGAATCCGTTAACCCCGCCAATCTGCA    60

Qy    21  PheAlaGlnSerGluProGluLeuLeuLysGluLeuSerValValIleValSerArgHisGly    40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188  TTCGCTCAGAGTGAGCCGGAGCTGGAAGTGTGGTGATTGTCAGTCGTCATGGT   247

Qy    41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248  TTCGCTCAGAGTGAGCCGGAGCTGGAAGTGTGGTGATTGTCAGTCGTCATGGT   307

Qy    41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308  GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA   367

Qy    61  ThrTrpProValLysLeuGlyLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu    80
          |||||||||||||||||||||||||||||||||||||||||||||
Db   308  GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA   367
```

RESULT 5
US-09-259-214-1
Sequence 1, Application US/09259214A
Patent No. 6110719
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: DIVR13170-1
; CURRENT APPLICATION NUMBER: US/09/259,214A
; CURRENT FILING DATE: 1999-03-01
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-259-214-1

US-10-601-319-10 (1-432) x US-09-259-214-1 (1-1323)

Alignment Scores:
Pred. No.:
Score: 2171.00
Percent Similarity: 97.7%
Best Local Similarity: 97.7%
Query Match: 96.1%

Pred. No.:   4.45e-222

Length: 1323
Matches: 422
Conservative: 0
Mismatches: 10
Indels: 0
Gaps: 0

Qy   1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla   20
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAGCGATTCTTAATCCCATTTTTATCTCTTCGATCCGTAAACCCGCAAATCTGCA   60

Qy   21 PheLeuAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly   40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCGCTCAGAGTGAGCCCGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy   61 ThrTrpProValLysProLeuGlyLeuGlyGlyTyrLeuGlyLeuLeuIleAlaTyrLeu   80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ACCTGGCCGGTAAAACCGGTTCGGCTGACACCCGCNGGGGTGGCTATCTC 240

Qy   81 GlyHisTyrTrpArgGlnAsnArgLeuValAlaAspGlyLeuLeuProLysLysCysPro   100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGACATTACCACGCCAGCGTCTGGGATGCTGCTGGACGGAAGAAAGGCTGCCCG 300

Qy   101 GlnSerGlyGlnValAlaIleIleAlaAspGlyLeuLeuProLysLysCysProThrHisGlyGlu   120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGTCTGGTCAGGTCGCGATTATTGCTGACGGGCTGCCCG 360

Qy   121 AlaPheAlaAlaGlyTyrLeuAlaProAspAspAspThr   140
           |||||||||||||||||||||||||||||||||||||||||
Db   361 GCCTTCGCGCGCGGCTGGCACCGTGACTGGTCAAGGGCGA 420

Qy   141 SerSerProAspProLeuPheAsnProLeuLysThrGlyIleLeuHisThrGlnLeuAlaAspThr   160
           |||||||||||||||||||||||||||||||||||||||||||
Db   421 TCCAGTCCCGATCCGTTATTTAATCCCTTAAAAACTGGGATAACGCG 480

Qy   161 AsnValThrAspAlaIleLeuGluArgAlaGlyGlySerIleAlaAspPheThrGlyHis   180
           ||||||||||||||||||||||||||||||||||||||||||
Db   481 AACGTGACTGACGCGATCCTCAGCAGGCAGGAGGGGTCAATTGCTGACTTTACGGGCAT   540

; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; SOFTWARE: FastSEQ for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 4
; EARLIER FILING DATE: 1999-03-01
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 09/291,931
; CURRENT FILING DATE: 1999-05-25
; CURRENT APPLICATION NUMBER: US/09/318,528
; FILE REFERENCE: 09010/029003
; TITLE OF INVENTION: NOVEL PHYTASE
; APPLICANT: Kretz, Keith
; GENERAL INFORMATION:
; Patent No. 6183740
; Sequence 1, Application US/09318528
US-09-318-528-1

RESULT 6

; OTHER INFORMATION: n = A,T,C or G
; LOCATION: (1)...(1323)
; NAME/KEY: misc_feature
; LOCATION: (1)...(1320)
; NAME/KEY: CDS
; FEATURE:
; ORGANISM: Escherichia coli

US-09-318-528-1

Alignment Scores:
Pred. No.:
Score: 2171.00
                          4.45e-222   Length: 1323
Percent Similarity: 97.7%   Matches: 422
Best Local Similarity: 97.7%   Conservative: 0
Query Match: 96.1%   Mismatches: 10
DB: 3   Indels: 0
                                             Gaps: 0

US-10-601-319-10 (1-432) x US-09-318-528-1 (1-1323)

[The remainder of the page consists of protein/DNA sequence alignment rows (Qy/Db) with alignment match bars, numbered in increments corresponding to residue/base positions from 1 through 1296/1323.]

```
Db    781  GATTCACCACCAGTGGAACACCTTGCTAAGTTTGCATAAGCGCAAATTTTATTGCTACAA  840

Qy    281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla  300

Db    841  CGCACGCCAGAGGTTGCCCGCGCGGCGACACCCCGGTTAATTGGATTTGATCAATGGCAGCG  900

Qy    301  LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu  320

Db    901  TTGACGGCCCCATCCACCCGCAAAAACAGGGCGTATGGGTGTGACATTACCCACTTCAGTACTG  960

Qy    321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340

Db    961  TTTATTGCCGGACACGATACTAACCTGGCAAATCTCGGGGGCGCCCTAGAGCTCAACTGG  1020

Qy    341  ThrLeuProGlyGlyAlaGlnLeuProGlyProGlyGlyAlaGluLeuAlaValPheGluLeuArgTrp  360

Db   1021  ACGGTTCCCGGTCAGCCGGGATAACGCCGGTGGATGAACTGGCGGTCTCAACGGGCTGG  1080

Qy    361  ArgGlyLeuSerGlnAspAspAsnSerGlnValIleGlnValSerLeuValPheGlnThrLeuGln  380

Db   1081  CGTCGGCGCTAAGGCGGATAACAGCCACGTGGATTCAGGTTTCCGCTGGTCTTCCAGACTTTACAG  1140

Qy    381  GlnMetArgGlnAspProLeuSerLeuAsnThrProLeuSerLeuAlaGlyLeuMetCysCysSerLeuAlaGlyLeuThr  400

Db   1141  CAGATGCGTCAGGATCCGCTGTCATTAAATACGGCCCCGCTTATTAAATCCTCTAAAAAACTGGGCGTTTGCCAATGGATAAGCG  1200

Qy    401  LeuAlaGlyCysGlyLeuArgAsnAlaGlyGlyMetCysCysSerProLeuAlaCysSerLeu  420

Db   1201  CTGGCAGGATGCGAAGAGCGGAAATGCCAGGGATTGTTCGGTTGGCAGGTTTTACGGCAA  1260

Qy    421  IleValAsnGlyValAlaArgIleLeuProAlaCysSerSerLeu  432

Db   1261  ATCGTGAATGAAGCACGATACCCCGGCGTGCCAGTTTG  1296
```

Alignment Scores:
```
Pred. No.:
Score:                 4.45e-222
Score:                 2171.00              Length:          1323
Percent Similarity:    97.7%                Matches:         1323
Best Local Similarity: 97.7%                Conservative:    422
Query Match:           96.1%                Mismatches:      10
DB:                    3                    Indels:          0
                                            Gaps:            0
```

US-10-601-319-10 (1-432) x US-09-291-931-1 (1-1323)

```
Qy      1  MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla  20
```

```
Db      1  ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAAATCGCA  60

Qy     21  PheIleAlaGlySerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly  40

Db     61  TTCGGTCAGAGTGAGCCGGAGCTGAAAGCTGAAAGTGTGGTGATTGTCAGTCGTCANGGT  120

Qy     41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro  60

Db    121  GTGCGTGCTCCAACCAAGGCCACCCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA  180

Qy     61  ThrTrpProValLeuGlyLeuGlyLeuThrProArgGlyGlyGlyLeuIleAlaTyrLeu  80

Db    181  ACCTGGCCGGTTCTGGGTCTGGGTTGGGCTGACACCGCGGGATGAACTGGCCTATCTC  240

Qy     81  GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro  100

Db    241  GGACATTACCGGCGCCAGCGCCTCGTGGCTGACGGCCTGCTGCCGAAAAAGGGCTGCCCG  300

Qy    101  GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu  120

Db    301  CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAAACAGGCGAA  360

Qy    121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr  140

Db    361  GCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACACACCCAGGCAGATACG  420

Qy    141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla  160

Db    421  TCCAGTCCCGGATCCGTTATTTAATCCTCTAAAAACTGGGGTTTGCCAACTGGATAACGCG  480

Qy    161  AsnValThrAspAlaIleLeuGluGluLeuArgAlaGlyGlySerIleLeuAlaAspPheThrGlyHis  180

Db    481  AACGTGACTGACGCGGCAATCCTCAGCAGGGGCAGGGGAGGGGCTCAAATTGCTGACTTTACCGGGCAT  540

Qy    181  TyrGlnThrAlaPheArgGluValLeuAspAsnPheProGlnSerAsnLeuCys  200

Db    541  CGGCAAACGGCGTTTCGCGAACTGGACAACTGGCCGCAGTCTAATCAAACTTTGTGC  600

Qy    201  LeuLysArgGluTyrGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220

Db    601  CTTAAACGTGAGAAACAGAGACGAAAGCTGTTCAATTACCAGTCAACTTACCATCAGGAACTC  660

Qy    221  LysValSerAlaAspCysValSerLeuThrGlyValAlaValSerLeuAlaSerMetLeuThr  240

Db    661  AAGGTGAGCGCCCGACAAATGTCTCATTAACCGGTCGGTAAGCCTCGCCATCAAATGCTGACG  720

Qy    241  GluIlePheLeuLeuGlnGlnAlaGlnGlnMetProGluProGlyTrpGlyArgIleThr  260

Db    721  GAGATATTTCTCTTGCCACAAGCACACGGGAATCCGGAGCCCGGGGAAGGATCACC  780

Qy    261  AspSerHisGlnThrProAsnThrLeuLeuSerLeuHisAsnAlaGlnProAspLeuLeuGln  280

Db    781  GATTCACCACCAGTGGAACACCTTGCTAAGTTTGCATAAGCGCAAATTTTATTGCTACAA  840

Qy    281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla  300

Db    841  CGCACGCCAGAGGTTGCCCGCGCGGCGACACCCCGGTTAATTGGATTTGATCAATGGCAGCG  900

Qy    301  LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu  320

Db    901  TTGACGGCCCCATCCACCCGCAAAAACAGGGCGTATGGGTGTGACATTACCCACTTCAGTACTG  960

Qy    321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340

Db    961  TTTATTGCCGGACACGATACTAACCTGGCAAATCTCGGGGGCGCCCTAGAGCTCAACTGG  1020

Qy    341  ThrLeuProGlyGlyAlaGlnLeuProGlyProGlyGlyAlaGluLeuAlaValPheGluLeuArgTrp  360

Db   1021  ACGGTTCCCGGTCAGCCGGGATAACGCCGGTGGATGAACTGGCGGTCTCAACGGGCTGG  1080

Qy    361  ArgGlyLeuSerGlnAspAspAsnSerGlnValIleGlnValSerLeuValPheGlnThrLeuGln  380
```

RESULT 8

US-09-515-1

Sequence 1, Application US/09580515
Patent No. 6720014
GENERAL INFORMATION:
APPLICANT: Kretz, Keith
TITLE OF INVENTION: NOVEL PHYTASE
FILE REFERENCE: 09010/029003
CURRENT APPLICATION NUMBER: US/09/580,515
CURRENT FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: 08/910,798
PRIOR FILING DATE: 1997-08-13
PRIOR APPLICATION NUMBER: 09/259,214
PRIOR FILING DATE: 1999-03-01
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1323
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1320)
NAME/KEY: misc_feature
LOCATION: (1)...(1323)
OTHER INFORMATION: n = A,T,C or G
US-09-580-515-1

Alignment Scores:
Pred. No.:
Score: 2171.00
Percent Similarity: 97.7%
Best Local Similarity: 97.7%
Query Match: 96.1%
DB:

US-10-601-319-10 (1-432) x US-09-580-515-1 (1-1323)

                    Length: 1323
                    Matches: 422
                    Conservative: 0
                    Mismatches: 10
                    Indels: 0
                    Gaps: 0

                    4.45e-222

RESULT 9
US-09-866-379-1

```
; Sequence 1, Application US/0986379
; Patent No. 6855365
; GENERAL INFORMATION:
;   APPLICANT: DIVERSA CORPORATION
;   APPLICANT: SHORT, Jay
;   APPLICANT: KRETZ, Keith
;   APPLICANT: GRAY, Kevin
;   APPLICANT: BARTON, Nelson
;   APPLICANT: GARRETT, James
;   APPLICANT: O'DONOGHUE, Eileen
;   TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
;   FILE REFERENCE: DIVER1370-7
;   CURRENT APPLICATION NUMBER: US/09/866,379
;   CURRENT FILING DATE: 2001-05-24
;   PRIOR APPLICATION NUMBER: US 09/580,515
;   PRIOR FILING DATE: 2000-05-25
;   PRIOR APPLICATION NUMBER: US 09/318,528
;   PRIOR FILING DATE: 1999-05-25
;   PRIOR APPLICATION NUMBER: US 09/291,931
;   PRIOR FILING DATE: 1999-04-13
;   PRIOR APPLICATION NUMBER: US 09/259,214
;   PRIOR FILING DATE: 1999-03-01
;   PRIOR APPLICATION NUMBER: US 08/910,798
;   PRIOR FILING DATE: 1997-08-13
;   NUMBER OF SEQ ID NOS: 10
;   SOFTWARE: Patentin version 3.1
;
; SEQ ID NO 1
;   LENGTH: 1323
;   TYPE: DNA
;   ORGANISM: Escherichia coli
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1323)
;   OTHER INFORMATION: n is any nucleotide
;   NAME/KEY: CDS
;   LOCATION: (1)..(1323)
;   OTHER INFORMATION:

US-09-866-379-1

Alignment Scores:
Pred. No.:
Score:            4.45e-222
Percent Similarity:    2171.00
Best Local Similarity:  97.7%
Query Match:            96.1%
DB:                      3

                    Length:         1323
                    Matches:         422
                    Conservative:      0
                    Mismatches:       10
                    Indels:            0
                    Gaps:              0

US-10-601-319-10  (1-432) x US-09-866-379-1 (1-1323)
```

```
Qy        1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60

Qy       21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGT 120

Qy       41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 GTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA 180

Qy       61 ThrTrpProValLysLeuGlyLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 ACCTGGCCGGTAAAACTGGGTTTGACGCCACGCGGCGGTGAGCTGATTGCTTATCTC 240

Qy       81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 GGACATTACCAACGCCAGCGTCTGGTCGCGGACGGATTGCTGCCGAAAAAGGGCTGCCCG 300

Qy      101 GlnSerGlyHisValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 CAGTCTGGTCATGTCGCGATTATTGCTGATGTGGACGAGCGTACCCGTAAAACAGGCGAA 360
```

```
Qy      121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 GCCTTCGCCGCCGGTCTGGCGCCCGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420

Qy      141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAGCTGGATAACGCG 480

Qy      161 AsnValThrAspAlaIleLeuGluArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 AACGTGACTGACGCCATTCTTGAGCGGGCAGGCGGCGGCGGTTCAATTGCTGACTTTACCGGGCAT 540

Qy      181 TyrGlnThrAlaPheAsnGluLeuGluLeuValLeuAsnPheProGlnSerAsnLeuCys 200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 CGGCAAAACGGCGGTTTCCGCGAACTGGAACTGGTGCTTAATTTCCCGCAATCAAACTTGTGC 600

Qy      201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 CTTAAACGTGAAAAGCAGGACGAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTC 660

Qy      221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 AAGGTGAGCGCCGACGTCTCATTAACCGGTGCGGTAAGCCTCGCCATCAATGCTGACG 720

Qy      241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 GAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGCTGGGGAAGGATCACC 780

Qy      261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 GATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTATTTGCTTACAA 840

Qy      281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTGGATTTGATCAAAACAGCG 900

Qy      301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 TTGACGCCCCATCCACCGCAAAAACAGGCCTATGGTGTGACATTACCGACTTCAGTACTG 960

Qy      321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGGGGCATTGGAGCTCAACTGG 1020

Qy      341 ThrLeuProGlyGlnProGlnAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 ACGCTTCCCGGTCAGCCGCGATAACACGCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGG 1080

Qy      361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140

Qy      381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCCCCGGGAGAGGTGAAACTGACC 1200

Qy      401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCGCTGGCAGGTTTTACGCAA 1260

Qy      421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
            ||||||||||||||||||||||||||||||||||||
Db     1261 ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTG 1296
```

```
RESULT 10
US-09-715-477-2
; Sequence 2, Application US/09715477
; Patent No. 6841370
; GENERAL INFORMATION:
;   APPLICANT: Lei, Xingen
;   TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
;   FILE REFERENCE: 19603/4031
;   CURRENT APPLICATION NUMBER: US/09/715,477
```

; CURRENT FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1486
; TYPE: DNA
; ORGANISM: Escherichia coli
Query Match:
DB:

US-09-715-477-2

US-10-601-319-10 (1-432) x US-09-715-477-2 (1-1486)

Alignment Scores:
Pred. No.:
Score: 3e-221
Percent Similarity: 2164.00
Best Local Similarity: 97.5%
Query Match: 97.5%
DB: 95.8%
                        3

Length: 1486
Matches: 421
Conservative: 0
Mismatches: 11
Indels: 0
Gaps: 0

```
Qy      1 MetLysAlaIleIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db    188 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGCTAACCCCGCAATCTGCA 247

Qy     21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db    248 TTCGCTCAGAGTGAGCCGGAGCTGAAACTGGAAAGTGTGGTGATTGTCAGCCGTCATGGT 307

Qy     41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaLaTrpPro 60
Db    308 GTGCGTGCCCCAACCAAAGCCACGCAACTGATGCAGGATGTCACCCCAGACGCCATGGCA 367

Qy     61 ThrThrProValLysLeuGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db    368 ACCACGCCGGTAAAACTGGGTGAGCTGACACCACGCGGTGGCGAACTTATCGCCTATCTC 427

Qy     81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysGlyCysProPro 100
Db    428 GGACATTACCAACGCCAGCGCGTCTGGGTGGCCGACGGATTGCTGCCGAAAAAAGGCTCCCG 487

Qy    101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db    488 CAGTCCGGTCAGGTCGCCGGATTATTGTCGATGTGGACGAGCGTACCCGTAAAAACAGGCGAA 547

Qy    121 AlaPheAlaAlaGlyLeuAlaProAspAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db    548 GCCTTCGCCGCCGGCCTGGCACCTGACGTGCAATAACCGTACATACCCAGGCAGATACG 607

Qy    141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAspAnAla 160
Db    608 TCCAGTCCGGACCCGTTATTTAATTCTCTAAAAACTGGCGTTTGCCAGCTGGATGACGCG 667

Qy    161 AsnValThrAspAlaIleLeuArgAlaAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db    668 AACGTGACTGACGCGATCCTGCGAGCGGCGGGTCAATGCTGACTTTACGGGGCAT 727

Qy    181 TyrGlnThrAlaPheArgGluLeuGluAsnPheProGlnSerAsnLeuCys 200
Db    728 CGGCAAACGGGCGTTTCGCGAACTGGAACGGGCTGCCTTAATTTCCGCAATCAAACTTGTGC 787

Qy    201 LeuLysArgGlyLysProGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db    788 CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCATTAACGCCAGGCCATTACCATCGGAACTC 847

Qy    221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db    848 AAGGTGAGCGCCGACAATGTTCATTAACGCGTGAGCCTCGCGATCAAATGCTGACG 907

Qy    241 GluIlePheLeuLeuGlnAsnGlnAlaGlyIleGlyMetProGlyLYsTrpGlyArgIleThr 260
Db    908 GAAATATTTCCTGCAACAACAGCAGGGAATGCCGGGACGCCGGGTGGGGAAGGGATCACT 967
```

RESULT 11
US-09-540-149A-9
; Sequence 9, Application US/0954014 9A
; Patent No. 6511699
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2791
; CURRENT APPLICATION NUMBER: US/09/540,149A
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 60/127,032
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1489
; TYPE: DNA
; ORGANISM: Escherichia coli
Query Match:
DB:

US-09-540-149A-9

US-10-601-319-10 (1-432) x US-09-540-149A-9 (1-1489)

Alignment Scores:
Pred. No.:
Score: 1.98e-219
Percent Similarity: 2147.00
Best Local Similarity: 96.8%
Query Match: 96.8%
DB: 95.1%
                        3

Length: 1489
Matches: 418
Conservative: 0
Mismatches: 14
Indels: 0
Gaps: 0

```
Qy      1 MetLysAlaIleIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db    182 ATGAAAGGGATCTTAATCCCATTTTTATCTCTTCTTTTGATTCCGCTAACCCCGCAATCTGCA 241

Qy     21 PheAlaGlnSerGluProGluLeuLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db    242 TTCGCTCAGAGTGAGCCGGAGCTGAAACTGGAAAGTGTGGTGATTGTCAGCCGTCATGGT 301
```

(continued in facing column)

```
Qy    261 AspSerHisGlnTrpAsnThrLeuLeuHisAsnAlaGlnPheAspLeuLeuGln 280
Db    968 GATTCACACCAGTGGAACACACTTGCTAAGTTTGCATAAGGCGCAATTTATTACTACAA 1027

Qy    281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
Db   1028 CGCACGCGCCAGAGGTTGCCCGCGCCACCCCGTTATTGGATTTGATCAAGACAGCG 1087

Qy    301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db   1088 TTGACGCCCCATCACCGCAAAAACAGGCGTATGGTGACAATTACCCACTTCAGTGCTG 1147

Qy    321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db   1148 TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGG 1207

Qy    341 GlnMetArgAspProLeuSerLeuAsnThrProProGlyGlyGluValLysLeuThr 400
Db   1208 CAGATGCGTGATCCATTAAGCCTGAATACGCCGCCGGGCGGTGAAGTGAAACTGACC 1267

Qy    401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db   1328 CTGGCAGGATGTGAAGAGCGAAATGCCCAGCAGGGCATGTGTTCGTTGGCCGGTTTTACGCAA 1387

Qy    421 IleValAsnGluIleAlaArgIleProAlaCysSerLeu 432
Db   1448 ATCGTGAATGAAGCGCCCATACCGGCCTGCAGTTTG 1483
```

RESULT 12
US-10-266-041A-9
; Sequence 9, Application US/10266041A
; Patent No. 6974690
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: PHOSPHATASESS WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2798
; CURRENT APPLICATION NUMBER: US/10/266,041A
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: 60/127,032
; PRIOR APPLICATION FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 09/540,149
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 1489
;   TYPE: DNA
;   ORGANISM: Escherichia coli
US-10-266-041A-9

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:

Pred. No.:	1.98e-219	Length:	1489
Score:	2147.00	Matches:	418
Percent Similarity:	96.8%	Conservative:	0
Best Local Similarity:	96.8%	Mismatches:	14
Query Match:	95.1%	Indels:	0
	4	Gaps:	0

US-10-601-319-10 (1-432) x US-10-266-041A-9 (1-1489)

Qy   1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   182 ATGAAAGGCGATTCTTAATCCCATTTTTATCTCTTTTGAATTCCGTTAACCCCGCAATCTGCA 241

Qy   21 PheLeuAlaGlnSerGluProGluLeuLysGlnLeuGlySerValValIleValSerArgHisGly 40
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   242 TTCGCTCAGAGTGAGCCCGAGCTGAAGCTGGGAAAAGTGTGGGTAGCATAAGTCATGGT 301

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   302 GTGCGTGCCCCAACCAAGGCCACCCAACTGATGCAAGACGTCACCCCAGACGCATGGCCA 361

Qy   61 ThrTrpProValLysGluGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   362 ACCTGGCCGGTAAAACTGGGTTGGCTGACACCACGCGTGGGTGAGCTAATCGCCTATCTC 421

Qy   81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   422 GGACATTACCAACCCCAGGGTCTCGGTGGCCGACGGATTGCTGCCGAAAAAGGGCTGCCG 481

Qy   101 GlnSerGlyGlnValAlaIleIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   482 CAGCCTGGTCAGGTCGCGATTATTGCTGACGTCGACGAGCGTACCCGTAAAACAGGCGAA 541

Qy   121 AlaPheAlaAlaGlyLeuValAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   542 GCCTTCGCCGCCGGTCTGGTCGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 601

Qy   141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   602 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCG 661

Qy   161 AsnValThrAspAlaIleLeuGluArgAlaGlyGlySerIleIleAlaAspPheThrGlyHis 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   662 AACGTGACTGACGCGATCCTCAGCAGGGCAGGAGGTCAAATTGCTGACTTTACCGGGCAT 721

RESULT 13
US-09-715-477-4
; Sequence 4, Application US/09715477
; Patent No. 6841370
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
; FILE REFERENCE: 1963/4031
; CURRENT APPLICATION NUMBER: US/09/715,477
; CURRENT FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 4
; LENGTH: 1486
; TYPE: DNA
; ORGANISM: Escherichia coli

```
US-09-715-477-4

US-10-601-319-10 (1-432) x US-09-715-477-4 (1-1486)

Alignment Scores:
Pred. No.:
Score:                    2144.00      Length:         1486
Percent Similarity:       97.2%        Matches:         418
                          96.8%        Conservative:      2
Best Local Similarity:    95.0%        Mismatches:       12
Query Match:                 3         Indels:            0
DB:                                    Gaps:              0
```

```
Qy    1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla   20
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  188 ATGAAAGCGATCTTAATCCCATTTTTAATCTCTTCTGATTCCCGTTAAACCCGCAATCTCGCA 247

Qy   21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly   40
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  248 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGCCGTCATGGT   307

Qy   41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro   60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  308 GTGCGTGCCCCAACCAAGGCAACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA   367

Qy   61 ThrTrpProValLysLeuGlyLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu     80
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  368 ACCTGGCCGGTGAAACTGGGTTTGACGCCTCGCGGTGGTGAGCTAATCGCCTATCTC     427

Qy   81 GlyHisTyrTrpArgGlnAsnGlyGlyValCysGlyLysCysPro                100
        |||||||||||||||||||||||||||||||||||||||||||||
Db  428 GGACATTACTGGCGTCAGAACGGCGGCGTCTGTGGGAAATGCCCG                487

Qy  101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  488 CAGTCCGGTCAGGTCGCGGATTATTGTCGATGTCGACGAGCGTACCCGTAAAACAGGCGAA 547

Qy  121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr  140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  548 GCCTTCGCCGCGGGTCTGGCCCCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG  607

Qy  141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla  160
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  608 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACGGGCGTTTGCCAACTGGATAACGCG  667

Qy  161 AsnValThrAspAlaIleLeuGluArgAlaGlyGlySerIleAlaAspPhePheGlyHis  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  668 AACGTGACTGACGCGATCCTCGAGCGGGCAGGGGGTTCAATTGCTGACTTTACCGGGCAT  727

Qy  181 TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys  200
        ||||||||||||||||||||||||||||||||||||||||||||||||::::|||||||||
Db  728 CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGAAC  787

Qy  201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  788 CTTAAACGTGAGAAAACAGAATCAAGCTGTAACTTAACGCAGGCATTACCATCGGAACTG  847

Qy  221 LysValSerAlaAspCysValSerLeuThrGlyAlaValAlaValSerLeuAlaSerMetLeuThr 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  848 AAGGTGAGCGCCGACAATGTTCATTAACCGGTGCGGTAAGCCTCGCCATCAATGCTGACG  907

Qy  241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGlyTyrTrpGlyTyrArgIleThr     260
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  908 GAAATATTTCTCCTGCAACAAGCACAGGGAATGCCGGGTTGGGGATACACT           967

Qy  261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln  280
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  968 GATTCACACCAGTGGAACACCTTGCTAAGTTGCATAACGCCCAATTTATTTACTACAA    1027

Qy  281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1028 CGCACGCCAGAGGTTGCCCGCAGCCGGGCCACCCCGTTATTGATGGAATTAAAGACAGCG 1087

Qy  301 LeuThrProHisProProGlyLysGlnAlaTyrGlyValThrLeuProThrSerValLeu  320
Db 1088 ...
```

```
Db  662 AACGTGGACTGACGCGATCCTCAGCAGGGCAGGAGGGTCAATTGCTGACTTTACGGGGCAT 721
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  181 TyrGlnThrAlaPheArgGluLeuLeuArgValLeuAsnPheProGlnSerAsnLeuCys  200

Db  722 CGGCAAACGGGCGTTTCGCAACGGGGTGCTTAATTTTTCCCAATTAAACTTGTGC       781
Qy  201 LeuLysArgGlyLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu  220

Db  782 CTTAACGTGAGAAACAGGAGGAAAACCTGTTCATTACGGCAGTACTCATCGGAGAACTC   841
Qy  221 LysValSerAlaAspCysValSerLeuThrGlyAlaValAlaValSerLeuAlaSerMetLeuThr 240

Db  842 AAGGTGAGCGCCGACAATGTTCATTAACCGGTGCGGTAAGCCTCGCCATCAATGCTGACG  901
Qy  241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGlyTyrTrpGlyTyrArgIleThr     260

Db  902 GAAATATTTCTCCTGCAACAAGCACAGGGAATGCCGGGTTGGGGAGGATACACT        961
Qy  261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln  280

Db  962 GATTCACACCAGTGGAACACCTTGCTAAGTTGCATAACGCCCAATTTATTTACTACAA   1021
Qy  281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla  300

Db 1022 CGCACGCCAGAGGTTGCCCGCAGTCGGCGCCACCCCGTTATTGGATTTGATCAAGACAGCG 1081
Qy  301 LeuThrProHisProProGlyLysGlnAlaTyrGlyValThrLeuProThrSerValLeu  320

Db 1082 TTGACGCCCCATCCACCCGGCAAAAACAGGCGTATGGCGTGACATTACCCACTTCAGTGCTG 1141
Qy  321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp  340

Db 1142 TTTATTGCCGGACACGACACTAATCTGGCAAATCTCGGCGGCGCACTGGAACTCAACTGG 1201
Qy  341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp  360

Db 1202 ACGCTTCCAGGTCAGCCGGATAACACGCCGCCGGGTGGCGAACTGGTGTTTGAACGCTGG 1261
Qy  361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln  380

Db 1262 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTGTCGCTGGTCTTCCAGACTTTACAG 1321
Qy  381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr  400

Db 1322 CAGATGCGTGATAAAACGCCCCTATCATTAAATACCCCGCCGGGAGGGGTGAAAACTGACC 1381
Qy  401 LeuAlaGlyLysCysGluGluValLysGlyAsnAlaLaGlnGlyMetCysSerLeuPheGlyThrGln 420

Db 1382 CTGGCAGGAAGTGTGAAGGAAGATGGCAGGGCAATGTGTTGCGTTTACGGCA        1441
Qy  421 IleValAlaAsnGluValAlaArgGlyIleProAlaCysSerLeu                432

Db 1442 ATCGTGGAAATGAAGCGCGCCCATACCGGCCGGCAGTTTG                    1477
```

Alignment Scores:

```
RESULT 14
US-08-910-798-1
Sequence 1, Application US/08910798
Patent No. 5876997
GENERAL INFORMATION:
  APPLICANT: KRETZ
  TITLE OF INVENTION: NOVEL PHYTASE
  NUMBER OF SEQUENCES: 4
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Fish & Richardson P.C.
    STREET: 4225 Executive Square, Suite 1400
    CITY: La Jolla
    STATE: California
    COUNTRY: US
    ZIP: 92037
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patentin Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/910,798
    FILING DATE: August 13, 1997
    CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
    NAME: HAILE, PH.D., LISA A.
    REGISTRATION NUMBER: 38,347
    REFERENCE/DOCKET NUMBER: 09010/029001
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 619/678-5070
    TELEFAX: 619/678-5099
  INFORMATION FOR SEQ ID NO: 1:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 1272 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
    IMMEDIATE SOURCE:
      CLONE: PHYTASE
    FEATURE:
      NAME/KEY:
      LOCATION:
; US-08-910-798-1
```

```
Pred. No.:                1.31e-210      Length:        1272
Score:                    2063.50        Matches:       405
Percent Similarity:       93.8%          Conservative:  0
Best Local Similarity:    93.8%          Mismatches:    10
Query Match:              91.4%          Indels:        17
DB:                       2              Gaps:          1
```

US-10-601-319-10 (1-432) x US-08-910-798-1 (1-1272)

```
Qy     1  MetLysAlaIleLeuIleLeuProPheLeuSerLeuLeuIleLeuProLeuThrProGlnSerAla   20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCTGCCGCCAACCTGCA   60

Qy    21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly   40
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  TTCGCTCAGAGTGAGCCGGAGCTGAAAGCTGAAAGTGTGGTGATTGTCGTCCGTCATGGT   120

Qy    41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GTGCGTGCTCCAACCAAGGCCACCCAACTGATGCAGGATGTCACCCCAGACGCATGGCCA   180

Qy    61  ThrTrpProValLeuGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu   80
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  ACCTGGCCGGTAAAACTGGGTGAGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC   240

Qy    81  GlyHisTyrTrpArgGlnAsnGlyLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro   100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  GGACATTACCAACGCCAGAATGGTCTCGTAGCCGACGGATTGCTGCCGAAAAAGGGCTGCCCG   300

Qy   101  GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu   120
          ||||||||||||||||||||||||||||||||||||||||||||
Db   301  CAGTCTGGTCAGGTTGCGATTATCGCCGATGTTGATGAGCGTACCCGTAAAACCGGC---   357

Qy   121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr   140
                                                      ||||||||||||
Db   358  --------------------------------------------CAGGCAGATACG   369

Qy   141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla   160
          ||||||||||||||||    ||||||||||||||||||||||||||||||||||||||
Db   370  TCCAGTCCCGGATCCGTTATTTAATCCTCTAAAAACTGGGGTTTGCCAGCTGGATAACGCG   429

Qy   161  AsnValThrAspAlaIleLeuGluGluValArgAlaLeuAlaAspValAspPheThrGlyHis   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   430  AACGTGACTGACGCGGATCCTCAGCAGGGCCAGAGGGGTCAAATTGCTGACTTTACCGGGCAT   489

Qy   181  TyrGlnInThrAlaPheArgGluLeuGlnArgValLeuAsnPheProGlnSerAsnLeuCys   200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   490  CGGCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAAYTTCCCGCAATCAAACTTGTGC   549

Qy   201  LeuLysArgGlyLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu   220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   550  CTTAAACGTGGAGAAACAGGACCAAAGCTGTTCATTAACGCAAGCCATTACCGACGAACTC   609

Qy   221  LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr   240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   610  AAGGTGAGCGCCGACATGTCTCATTAACGGTTGCGGTAAGCCTCGCAATCAATGCTCACG   669

Qy   241  GluIlePheLeuLeuGlnInGlyAspGlyMetProGlyTyrGlyArgIleThr   260
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   670  GAGATATTTCTCCTGCAACAAGGACGAGGGATGCCGGGCTATGGGAGGAGGATCACC   729

Qy   261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln   280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   730  GATTCACACAGTGGAACACCTTGCTAAGTTTGCATAAGCGCCAAATTTTATTTGCTACAA   789

Qy   281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla   300
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   790  CGCACGCCAGAGGTTGCCCGCAGCCAGCGCGCCACCCCGTTATTGGATCATTAAAACAGCG   849

Qy   301  LeuThrProHisProProGlyLeuAlaTyrGlyValThrLeuProThrSerValLeu   320
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   850  TTGACGCCCCATCCACCCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTACTG   909
```

```
Qy   321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp   340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1148  TTTATTGCCGGACACGATACTAATCTGGCAAATCTCGGCGGGGCCGGAGCTCAACTGG   1207

Qy   341  ThrLeuProGlyGlnProLeuAspAsnThrProProGlyValGlyValPheGluArgTrp   360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1208  ACGCTTCCAGGTCAGCCGCGATAACGACGCCGCCAGGTGGTGTGAACGGCTGG   1267

Qy   361  ArgArgLeuSerAspAsnAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln   380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1268  CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG   1327

Qy   381  GlnMetArgAspGlyLysThrProLeuSerLeuAsnThrProProGlyValGlyValLysLeuThr   400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1328  CAGATGCGTGATGGAAAAACGCCGCTATCATTAAATACGCCCCCGGAGAGGTGAAAACTGACC   1387

Qy   401  LeuAlaGlyTyrGluGlyLeuGlyArgAsnAlaGlnGlyMetCysSerSerLeuAlaGlyPheThrGln   420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1388  CTGGCAGGATTGTGAAGGACGGAAATGCGCAGGGCATGTGTTCGTTCGTTGGCCGGTTTTACGGCAA   1447

Qy   421  IleValAsnGluAlaAlaArgIleProAlaCysSerLeu   432
          |||||||||||||||||||||||||||||||||||||||
Db  1448  ATCGTGAATGAAGCGGCCCGATACCGGGCGTGCAGTTTG   1483
```

RESULT 15
US-09-489-039A-341

; Sequence 341, Application US/0948903 9A
; Patent No. 6610836
; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 341
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-341

US-10-601-319-10 (1-432) x US-09-489-039A-341 (1-1266)

Alignment Scores:
Pred. No.:
Score: 543.00
Percent Similarity: 49.4%
Best Local Similarity: 32.6%
Query Match: 24.0%
DB: 3

Length: 1266
Matches: 136
Conservative: 70
Mismatches: 181
Indels: 30
Gaps: 10

4.93e-48

Qy 321 PheIleAlaGlyTyrHisAspThrAsnLeuAlaAlaAsnLeuGlyTyrGlyAlaLeuGlyLeuGlyLeuAsnTrp 340
        ||||||| |||| ||:::|| |||||| | |||| ||||::||| ||||||||||||||| ||||||
Db 910 TTTATTGCCGGACACGATACTAATCTGCAAAATCTGGCAAAATCTGGCAACTGGGCCGCGCACTGGGACTCAACTGG 969

Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
        |||||| :::|||: ||| ||||| ||||| ||||||||||:||||| |||
Db 970 ACGCTTCCCGGTCAGCCGGATAACACGCCCCCAGGTGGTGAACTGGGTTGAACGCTG 1029

Qy 361 ArgArgLeuSerAsnAspAsnSerGlnIleThrProIleGlnValSerLeuValPheGlnHisThrLeuGln 380
        |||| ::: || ||| | |||||| || ||||||||||||:|| || || |||||| ||||:|:
Db 1030 CGTTCGGCTAAGCGATAACAGCCAGTGGGATTCAGGTTTCGCTGGTTCTTCCAGACTTTACAG 1089

Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
        ||||| ||||||||||||||::|||: |||||||||:||:|| ||||||| |||
Db 1090 CAGATGCGTGATAAAACGCCCGCTGTCATTAAAATACCGCCCGGCGAGGGTGAAACTGACC 1149

Qy 401 LeuAlaGlyCysGluGluValArgAsnAlaGlyLysMetCysSerSerLeuAlaGlyPheThrGln 420
        ||||| |||| ||||| ::::|| ||||| |||| :||||| ||||| | ||||||| |
Db 1150 CTGGCAGGATGTGAAGAGGTGAGAAATGGGCAGGGCATGTGTTCGTCGTTGGCAGGTTTTACGGCAA 1209

Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
        |||:|:::||||||||||||||| ||| | |:
Db 1210 ATCGTGAAATGAAGCACGCCATTACCGGGCTGCAGTTG 1245

Qy 113 GluArgThrArgLysThrGlyGluAlaPheAlaAlaGlyLeuAlaProAspCysAlaIle 132
        ||:::||: ||||||| ||| :||||||||:::|||||| ||||| ||||||||:
Db 367 CAGCGCACCGTCGCCACCGACCCGGAGTTCTTTATCACCAGATTCCCGGGCGATC 426

Qy 133 ThrValHisThrGlnAlaAspThrSerProAspProLeuPheAsnProLeuLysThr 152
        ||||| ||||| || ::: |||||| | :::|||||| ||||| ||
Db 427 CCGGTGCATCACCAGCTCAGATGGGCACCATGGACCCGGACCTTCAACCCGGCTTATTACC 486

Qy 153 GlyValCysGlnLeuAspAsnAlaAsnValThrArgAlaIle---LeuGlnArgAlaGly 171
        ||| | |||||:::|| |||||| ||:::|||||| ||:::||||| |||||||
Db 487 GACGACTCCCGGCGGTTTCGCGAGAAAGGCTGCGATGGGAAAAAGAGCCCCAGGA 546

Qy 172 GlySerIleAlaAspPheThrGlyHisTyrGlnThrArgAlaPheArgGluGluArgVal 191
        || :|| | |||| |||||||| ||||||| ||| |||||||| ||||| |:
Db 547 AGGCAGTTGACCGAG---------------------AGCTACAAGCTGCTGGAGACGATG 585

Qy 192 LeuAsnPheProGlnSerAsnLeuCysLeuLysArgGlyGlnAspGluSerCysSer 211
        ||||| ||||| | |||||| |||||| |||||:::||| |:||:
Db 586 ATTGAACTACCGCAACTCGCCCTCCTGC-----------AAAGAAGAAGGCTTCGTTCC 633

Qy 212 LeuThrGlnAlaLeuProSerGluLeuLysValSerAlaAspCysValSerLeuThrGly 231
        || |||||| |||||| ||||||||||:::|||||||||| |:::
Db 634 CTGACGCAGGCGTAAAGATACCTTTAGCCGGGTTATCAGCAGGAGCCCGGGAGTGTCCCGA 693

Qy 232 AlaValSerLeuAlaSerMetLeuThrGluIleIlePheLeuLeuGlnIleAlaGlnGlyMet 251
        ||||||:|||||||| | ||| | |||| ::|||| |||||| || |||:
Db 694 CCGGTGAAGGTGGCAACTCGGTGAGACGCGTTCACCCTGCAAATATTATGAAGGTTC 753

Qy 252 Pro-------GluProGlyTyrTrpGlyArgIleThrAspSerHisGlnThrProAsnThrLeuLeu 269
        ||| |||||||| |||| :::::||||:::|:||||:||
Db 754 CCGAAAGATCAGGTGGCGTGGGGGGAGATCCGCCAGCAATAAGCAGTGCGGACTGCTGTCG 813

Qy 270 SerLeuHisAsnAlaGlnPheAspLeuLeuGlnArgThrProGluAlaLeuArgSerArg 289
        ||| |||||||||| | || |:::||| | ||||||||||||:::
Db 814 AAGCTGAAAAATGGCTATCAGGATAGTCTGTTTACCTCGGTGGCCGCAAAACGTC 873

Qy 290 AlaThrProLeuLeuAsnProLeuIleLeuLysThrAlaLeuThrProHisProProGlnLysGln 309
        ||||||||||| ||||| |:::|| | | |:::|||||||:||:||||
Db 874 GCCAAACCGCTGGTTAAATATATCGATAACGGCCTGGGCCAGCAAG-- 930

Qy 310 AlaTyrGlyValThrLeuProThrSerValLeuPheIleLeuAlaGlyHisHisAspThrAsnLeu 329
        ||:::||| |||| | ::::|||||||| | ||||||| |||:::
Db 931 -------------GCGAAAGTCACGCGCTGCGTCGGTGGGACACGACTCCGAATATC 969

Qy 330 AlaAsnLeuGlyAlaAlaLeuGlyLeuAsn---TrpThrLeuProGlyGlnProAspAsn 348
        ||||| ||||| ||||||||::: ||| ||||||||||||:::
Db 970 GCTTCGGCTGCTCACCGCCTACAACTGCCCGCCAGTATGACCGCC 1029

Qy 349 ThrProProGlyGlyGluLeuValPheGluArgTrpArgArgLeuSerAsnAspAsnSerGln 368
        ||||||||||| |||| | ||||| ::: :||||||||||||:::
Db 1030 ACTCCTATCGGCGGCGAAACTGCTGTTTCAGCGCATGACGAGCGCCGCGGCGCAATCGGGAT 1089

Qy 369 TrpIleGlnValSerLeuValPheGlnHisGlnMetArgAspLysThrProLeu 388
        |||::: ||| | ||||||:::|||||||||||:::
Db 1090 CTGATGAAGATGAGTAGTCTATCAAAGTACCGGCAACTGCGTTA 1149

Qy 389 SerLeuAsnThrProProGlyGluValLysLeuThrLeuAlaGlyCysGluGluValArgAsn 408
        |||||:::||||||||||||||:|:::||||| |::: :|||||||||:::
Db 1150 ACCCTGCAGGCCGCGCCCGCAGCGGGTCACTCTCGGCGCTTAACGGCGTCCGGTT---GAT 1206

Qy 409 AlaGlyLysMetCysSerSerLeuAlaGlyPheThrGlnIleValAsnGluAla 425
        |||||| | :::||||::: |||| | ||||||||||:::
Db 1207 GACCAGGGGATTCTGCCCGCCTGGAGACCTTTAAAATGGTGATCAACGAGGCG 1257

OM protein - protein search, using sw model

                    GenCore version 5.1.9
                    Copyright (c) 1993 - 2006 Bioacceleration Ltd.

Run on: June 13, 2006, 10:45:37 ; Search time 140.697 Seconds
                                               (without alignments)
                                               1422.265 Million cell updates/sec

Title: US-10-601-319-10
Perfect score: 2258
Sequence: 1 MKAILIPFLSLIIPLTPQSA..........CSLAGFTQIVNBARIPACSL 432

Scoring table: BLOSUM62
                    Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 46321485B residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published Applications AA Main:*
                    1: /EMC_Celerra_SID53/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
                    2: /EMC_Celerra_SID53/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
                    3: /EMC_Celerra_SID53/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
                    4: /EMC_Celerra_SID53/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
                    5: /EMC_Celerra_SID53/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
                    6: /EMC_Celerra_SID53/ptodata/2/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2258	100.0	432	4	US-10-601-319-10	Sequence 10, Appl
2	2258	100.0	432	5	US-10-933-115-10	Sequence 10, Appl
3	2258	100.0	432	6	US-11-056-354-2	Sequence 2, Appl
4	2258	100.0	436	4	US-10-156-660-2	Sequence 2, Appl
5	2228	98.7	430	3	US-09-866-379-10	Sequence 10, Appl
6	2228	98.7	432	3	US-09-866-379-8	Sequence 8, Appl
7	2182	96.6	432	4	US-10-156-660-4	Sequence 4, Appl
8	2182	96.6	432	4	US-10-282-122A-43351	Sequence 43351, A
9	2182	96.6	432	4	US-10-601-319-8	Sequence 8, Appl
10	2182	96.6	432	5	US-10-472-317-41	Sequence 41, Appl
11	2182	96.6	432	5	US-10-933-115-8	Sequence 8, Appl
12	2182	96.6	432	6	US-11-056-354-4	Sequence 4, Appl
13	2177	96.4	432	6	US-11-018-709-1	Sequence 1, Appl
14	2177	96.1	440	3	US-09-777-566A-2	Sequence 2, Appl
15	2171	96.1	440	3	US-09-866-379-2	Sequence 2, Appl
16	2171	96.1	440	4	US-10-034-985-2	Sequence 2, Appl
17	2171	96.1	440	4	US-10-430-356-2	Sequence 2, Appl
18	2171	96.1	440	5	US-10-933-115-2	Sequence 2, Appl
19	2171	96.1	440	5	US-10-933-115-2	Sequence 2, Appl
20	2168.5	96.0	437	6	US-11-074-522-12	Sequence 12, Appl
21	2157	95.5	412	4	US-10-334-672-1	Sequence 1, Appl
22	2157	95.5	412	4	US-10-334-671-1	Sequence 1, Appl
23	2157	95.5	432	6	US-11-018-709-3	Sequence 3, Appl
24	2153	95.3	410	4	US-10-021-723A-13	Sequence 13, Appl
25	2153	95.3	410	4	US-10-021-723A-15	Sequence 15, Appl
26	2147	95.1	432	4	US-10-284-962-3	Sequence 3, Appl
27	2147	95.1	432	4	US-10-284-962-14	Sequence 14, Appl
28	2147	95.1	433	4	US-10-266-041-1	Sequence 1, Appl
29	2144	95.0	432	5	US-10-284-962-5	Sequence 5, Appl
30	2112	93.5	432	6	US-10-021-723A-16	Sequence 16, Appl
31	1192	52.8	261	5	US-10-450-763-56057	Sequence 56057, A
32	976.5	43.2	441	4	US-10-021-723A-12	Sequence 12, Appl
33	959.5	42.5	441	4	US-10-282-122A-77792	Sequence 77792, A
34	951.5	42.1	441	4	US-10-021-723A-2	Sequence 2, Appl
35	948.5	42.0	420	4	US-10-021-723A-4	Sequence 4, Appl
36	782	34.6	476	4	US-10-021-723A-10	Sequence 10, Appl
37	637	28.2	144	5	US-10-450-763-54615	Sequence 54615, A
38	627.5	27.8	318	4	US-10-021-723A-14	Sequence 14, Appl
39	475.5	21.1	409	4	US-10-021-723A-8	Sequence 8, Appl
40	462.5	20.5	421	4	US-10-021-723A-6	Sequence 6, Appl
41	414	18.3	441	6	US-11-098-686-10683	Sequence 10683, A
42	310	13.7	118	5	US-10-450-763-56055	Sequence 56055, A
43	197	8.7	426	4	US-10-257-174-44	Sequence 44, Appl
44	197	8.7	426	4	US-10-343-357-6	Sequence 6, Appl
45	197	8.7	426	5	US-10-450-763-53707	Sequence 53707, A

ALIGNMENTS

RESULT 1
US-10-601-319-10
; Sequence 10, Application US/10601319
; Publication No. US2004009196BA1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING THEM
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified phytase enzyme
US-10-601-319-10

Query Match   100.0%; Score 2258; DB 4; Length 432;
Best Local Similarity   100.0%;
  Matches 432; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MKAILIPFLSLIIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKAILIPFLSLIIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP 60

Qy   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKKCGCPQSGQVAIIADVDERTRKTGE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKKCGCPQSGQVAIIADVDERTRKTGE 120

RESULT 2
US-10-933-115-10
; Sequence 10, Application US/10933115
; Publication No. US20050281792A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 564462001822
; CURRENT APPLICATION NUMBER: US/10/933,115
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified phytase enzyme
US-10-933-115-10

Query Match   100.0%; Score 2258; DB 5; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.8e-196;
Matches 432; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP 60

Qy 121 AFAAGLAPDCAITVHTQADTSSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 AFAAGLAPDCAITVHTQADTSSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180

Qy 181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240

Qy 241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300

Qy 301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNIANLGGALELNWTLPGQPDNTPPGGELVFERW 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNIANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy 361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy 421 IVNEARIPACSL 432
         ||||||||||||
Db 421 IVNEARIPACSL 432

RESULT 3
US-11-056-354-2
; Sequence 2, Application US/11056354
; Publication No. US20050246780A1
; GENERAL INFORMATION:
; APPLICANT: SHORT, Jay M.
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin A.
; APPLICANT: BARTON, Nelson R.
; APPLICANT: GARRETT, James B.
; APPLICANT: O' DONOGHUE, Eileen
; APPLICANT: MATHUR, Eric J.
; TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM AND METHODS
; TITLE OF INVENTION: FOR MAKING AND USING THEM
; FILE REFERENCE: 564462001803/D1370-9C1
; CURRENT APPLICATION NUMBER: US/11/056,354
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 10/156,660
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 2
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Modified phytase enzyme
; SOFTWARE: Patentin version 3.1
US-11-056-354-2

Query Match   100.0%; Score 2258; DB 5; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.8e-196;
Matches 432; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKGGCPQSGQVAIIADVDERTRKTGE 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKGGCPQSGQVAIIADVDERTRKTGE 120

Qy 121 AFAAGLAPDCAITVHTQADTSSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 AFAAGLAPDCAITVHTQADTSSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180

Qy 181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240

Qy 241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300

Qy 301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNIANLGGALELNWTLPGQPDNTPPGGELVFERW 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNIANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy 361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy 421 IVNEARIPACSL 432
         ||||||||||||
Db 421 IVNEARIPACSL 432

RESULT 4
US-10-156-660-2
; Sequence 2, Application US/10156660
; Publication No. US20030103958A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson R.
; APPLICANT: Garrett, James B.
; APPLICANT: O'Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM
; TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 09010-029007
; CURRENT APPLICATION NUMBER: US/10/156,660
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified phytase enzyme
US-10-156-660-2

```
Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60

Qy   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH  180

Qy  181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLLDLIKTA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLLDLIKTA 300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLLPGQPDNTPPGGELVFERW 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLLPGQPDNTPPGGELVFERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy  421 IVNEARIPACSL 432
        ||||||||||||
Db  421 IVNEARIPACSL 432
```

RESULT 5
US-09-866-379-10
; Sequence 10, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 10
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Modified phytase
US-09-866-379-10

```
Query Match             100.0%;  Score 2258;  DB 4;  Length 436;
Best Local Similarity   100.0%;  Pred. No. 1. 1.8e-196;
Matches  432;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60

Qy   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60

Qy  121 AFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120

Qy  181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH  180

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLLDLIKTA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLLPGQPDNTPPGGELVFERW 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLLDLIKTA 300

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLLPGQPDNTPPGGELVFERW 360

Qy  421 IVNEARIPACSL 432
        ||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

        IVNEARIPACSL 432
```

Query Match
Best Local Similarity   98.7%;   Score 2228; DB 3; Length 430;
                        99.5%;   Pred. No. 9.7e-194;
Matches 430; Conservative   0; Mismatches   0; Indels   2; Gaps   2;

```
Qy    1 MKAILIPFLSLLILPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLILPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP  60

Qy   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
        |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILB-AGGSIADFTGH 179

Qy  181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
        |||||||||||||||||||||||||||||||||||||||||||||  |||||||||||||
Db  180 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 239

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
        |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db  240 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFD-LQRTPEVARSRATPLLDLIKTA 298

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  299 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 358

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  359 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 418

Qy  421 IVNEARIPACSL 432
        ||||||||||||
Db  419 IVNEARIPACSL 430
```

RESULT 6
US-09-866-379-8
Sequence 8, Application US/09866379
Patent No. US20020136754A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: SHORT, Jay
APPLICANT: KRETZ, Keith
APPLICANT: GRAY, Kevin
APPLICANT: BARTON, Nelson
APPLICANT: GARRETT, James
APPLICANT: O'DONOGHUE, Eileen
TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
FILE REFERENCE: DIVER1370-7
CURRENT APPLICATION NUMBER: US/09/866,379
CURRENT FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 09/580,515
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentin version 3.1
SEQ ID NO 8
LENGTH: 432
TYPE: PRT
ORGANISM: Escherichia coli
US-09-866-379-8

Query Match
Best Local Similarity   96.6%;   Score 2182; DB 3; Length 432;
                        98.1%;   Pred. No. 1.5e-189;

Matches 424; Conservative   0; Mismatches   8; Indels   0; Gaps   0;

```
Qy    1 MKAILIPFLSLLILPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLILPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP  60

Qy   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
        |||||| |||||||||||||||| ||||||||||  ||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
        |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy  181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
         |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db  181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy  421 IVNEARIPACSL 432
        ||||||||||||
Db  421 IVNEARIPACSL 432
```

RESULT 7
US-10-156-660-4
Sequence 4, Application US/10156660
Publication No. US20030103958A1
GENERAL INFORMATION:
APPLICANT: Short, Jay M.
APPLICANT: Kretz, Keith
APPLICANT: Gray, Kevin A.
APPLICANT: Barton, Nelson R.
APPLICANT: Garrett, James B.
APPLICANT: O'Donoghue, Eileen
APPLICANT: Mathur, Eric J.
TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM AND METHODS FOR MAKING AND USING THEM
FILE REFERENCE: 09010-029007
CURRENT APPLICATION NUMBER: US/10/156,660
CURRENT FILING DATE: 2002-10-01
PRIOR APPLICATION NUMBER: US 09/866,379
PRIOR FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 09/580,515
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 4
LENGTH: 432
TYPE: PRT
ORGANISM: Escherichia coli
US-10-156-660-4

Query Match   96.6%;   Score 2182; DB 4; Length 432;

RESULT 8
US-10-282-122A-43351
Sequence 43351, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA_034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patentin version 3.1

SEQ ID NO 43351
LENGTH: 432
TYPE: PRT
ORGANISM: Escherichia coli
US-10-282-122A-43351

Query Match   98.6%;   Score 2182;   DB 4;   Length 432;
Best Local Similarity   98.1%;   Pred. No. 1.5e-189;
Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATOLMQDVTPDAWP 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATOLMQDVTPDAWP 60

Qy   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGELTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy   121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180

Qy   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy   241 EIFLLQOAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 EIFLLQOAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPCQPDNTPPGGELVFERW 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPCQPDNTPPGGELVFERW 360

Qy   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy   421 IVNEARIPACSL 432
          ||||||||||||
Db   421 IVNEARIPACSL 432

RESULT 9
US-10-601-319-8
Sequence 8, Application US/10601319
Publication No. US20040091968A1
GENERAL INFORMATION:

APPLICANT: Short, Jay M.
APPLICANT: Kretz, Keith A.
APPLICANT: Gray, Kevin A.
APPLICANT: Barton, Nelson Robert
APPLICANT: Garrett, James B.
APPLICANT: O' Donoghue, Eileen
APPLICANT: Mathur, Eric J.
TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
   AND USING THEM
FILE REFERENCE: 09010-029011
CURRENT APPLICATION NUMBER: US/10/601,319
CURRENT FILING DATE: 2003-06-20
PRIOR APPLICATION NUMBER: US 09/866,379
PRIOR FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 09/580,515
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25

Best Local Similarity   98.1%;   Pred. No. 1.5e-189;
Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATOLMQDVTPDAWP 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATOLMQDVTPDAWP 60

Qy   61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGELTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE 120

Qy   121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180

Qy   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy   241 EIFLLQOAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 EIFLLQOAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Qy   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPCQPDNTPPGGELVFERW 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPCQPDNTPPGGELVFERW 360

Qy   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy   421 IVNEARIPACSL 432
          ||||||||||||
Db   421 IVNEARIPACSL 432

```
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-601-319-8

  Query Match    96.6%;  Score 2182;  DB 4;  Length 432;
  Best Local Similarity  98.1%;  Pred. No. 1.5e-189;
  Matches  424;  Conservative  0;  Mismatches  8;  Indels  0;  Gaps  0;

Qy    1 MKAILIPFLSLLLIPLLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLLIPLLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYWRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYWRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH  180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH  180

Qy  181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240

Qy  241 EIFLLQQAQGMPEPGWGKRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGKRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA  300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW  360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW  360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420

Qy  421 IVNEARIPACSL 432
        ||||||||||||
Db  421 IVNEARIPACSL 432


RESULT 10
US-10-472-317-41
; Sequence 41, Application US/10472317
; Publication No. US20040018562A1
; GENERAL INFORMATION:
; APPLICANT: Cargill, Incorporated
; TITLE OF INVENTION: Myo-Inositol Oxygenases
; FILE REFERENCE: 10829/003USI
; CURRENT APPLICATION NUMBER: US/10/472,317
; CURRENT FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/US02/08404
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: US 60/277,148
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-472-317-41
```

```
  Query Match    96.6%;  Score 2182;  DB 4;  Length 432;
  Best Local Similarity  98.1%;  Pred. No. 1.5e-189;
  Matches  424;  Conservative  0;  Mismatches  8;  Indels  0;  Gaps  0;

Qy    1 MKAILIPFLSLLLIPLLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKAILIPFLSLLLIPLLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP  60

Qy   61 TWPVKLGWLTPRGGELIAYLGHYWRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TWPVKLGWLTPRGGELIAYLGHYWRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120

Qy  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH  180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH  180

Qy  181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240

Qy  241 EIFLLQQAQGMPEPGWGKRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EIFLLQQAQGMPEPGWGKRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA  300

Qy  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW  360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVFERW  360

Qy  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420

Qy  421 IVNEARIPACSL 432
        ||||||||||||
Db  421 IVNEARIPACSL 432


RESULT 11
US-10-933-115-8
; Sequence 8, Application US/10933115
; Publication No. US20050281792A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
;   THEREOF
; FILE REFERENCE: 564462001822
; CURRENT APPLICATION NUMBER: US/10/933,115
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO 8
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-933-115-8
```

```
Query Match        96.6%;   Score 2182;   DB 5;   Length 432;
Best Local Similarity   98.1%;   Pred. No. 1.5e-189;
Matches 424;   Conservative  0;   Mismatches  8;   Indels  0;   Gaps  0;

Qy    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60

Qy   61  TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120

Qy  121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180

Qy  181  RQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  RQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240

Qy  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA  300

Qy  301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360

Qy  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420

Qy  421  IVNEARIPACSL  432
         ||||||||||||
Db  421  IVNEARIPACSL  432
```

RESULT 12
US-11-056-354-4
```
; Sequence 4, Application No. US200502467801
; Publication No. US20050246780A1
; GENERAL INFORMATION:
; APPLICANT: SHORT, Jay M.
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin A.
; APPLICANT: BARTON, Nelson R.
; APPLICANT: GARRETT, James B.
; APPLICANT: O'DONOGHUE, Eileen
; APPLICANT: MATHUR, Eric J.
; TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM AND METHODS
; TITLE OF INVENTION: FOR MAKING AND USING THEM
; FILE REFERENCE: 564462001803/D1370-9C1
; CURRENT APPLICATION NUMBER: US/11/056,354
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 10/156,660
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.1
;
; SEQ ID NO 4
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli appA phytase
US-11-056-354-4
```

```
Query Match        96.6%;   Score 2182;   DB 6;   Length 432;
Best Local Similarity   98.1%;   Pred. No. 1.5e-189;
Matches 424;   Conservative  0;   Mismatches  8;   Indels  0;   Gaps  0;

Qy    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60

Qy   61  TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLAKKGCPQSGQVAIIADVDERTRKTGE  120

Qy  121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH  180

Qy  181  RQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  RQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240

Qy  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA  300

Qy  301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW  360

Qy  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420

Qy  421  IVNEARIPACSL  432
         ||||||||||||
Db  421  IVNEARIPACSL  432
```

RESULT 13
US-11-018-709-1
```
; Sequence 1, Application US/11018709
; Publication No. US20050095691A1
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
; FILE REFERENCE: 19603/4031
; CURRENT APPLICATION NUMBER: US/11/018,709
; CURRENT FILING DATE: 2004-12-20
; PRIOR APPLICATION NUMBER: US/09/715,477
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
;
; SEQ ID NO 1
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-11-018-709-1
```

```
Query Match        96.4%;   Score 2177;   DB 6;   Length 432;
Best Local Similarity   97.9%;   Pred. No. 4.3e-189;
Matches 423;   Conservative  0;   Mismatches  9;   Indels  0;   Gaps  0;

Qy    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60

Qy   61  TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWP   60
```

```
Qy      61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQPGQVAIIADVDERTRKTGE 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQPGQVAIIADVDERTRKTGE 180

Qy     121 AFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 RQTAFRELERVLNFPQQSNLCLKREKQDESCSLTQALPSELKYSADCVSLTGAVSLASMLT 240

Qy     181 YQTAFRELERVLNFPQQSNLCLKREKQDESCSLTQALPSELKYSADCVSLTGAVSLASMLT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLLDLLIKTA 300

Qy     241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLLDLLIKTA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVPERW 360

Qy     301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVPERW 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy     361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 IVNEARIPACSL 432

Qy     421 IVNEARIPACSL 432
           ||||||||||||
Db     421 IVNEARIPACSL 432
```

RESULT 14
US-09-777-566A-2

```
;  Sequence 2, Application US/09777566A
;  Patent No. US20010055788A1
;  GENERAL INFORMATION:
;     APPLICANT: DIVERSA CORPORATION
;     APPLICANT: SHORT, Jay
;     APPLICANT: KRETZ, Keith
;     TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
;     FILE REFERENCE: DIVER1370-6
;     CURRENT APPLICATION NUMBER: US/09/777,566A
;     CURRENT FILING DATE:
;     PRIOR APPLICATION NUMBER: US 09/318,528
;     PRIOR FILING DATE: 1999-05-25
;     PRIOR APPLICATION NUMBER: US 09/291,931
;     PRIOR FILING DATE: 1999-04-13
;     PRIOR APPLICATION NUMBER: US 09/259,214
;     PRIOR FILING DATE: 1999-03-01
;     PRIOR APPLICATION NUMBER: US 08/910,798
;     PRIOR FILING DATE: 1997-08-13
;     NUMBER OF SEQ ID NOS: 4
;     SOFTWARE: Patentin version 3.0
;  SEQ ID NO 2
;     LENGTH: 440
;     TYPE: PRT
;     ORGANISM: Escherichia coli
;     FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (1)..(1323)
;     OTHER INFORMATION: n is any nucleotide
US-09-777-566A-2
```

```
Query Match             96.1%;  Score 2171;  DB 3;  Length 440;
Best Local Similarity   97.7%;  Pred. No. 1.6e-188;
Matches  422;  Conservative  0;  Mismatches 10;  Indels  0;  Gaps  0;

Qy       1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP 60

Qy      61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQPGQVAIIADVDERTRKTGE 120

Qy     121 AFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
```

```
Db     121 AFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180

Qy     181 YQTAFRELERVLNFPQQSNLCLKREKQDESCSLTQALPSELKYSADCVSLTGAVSLASMLT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 RQTAFRELERVLNFPQQSNLCLKREKQDESCSLTQALPSELKYSADNVSLTGAVSLASMLT 240

Qy     241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLLDLLIKTA 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLLDLLIMAA 300

Qy     301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVPERW 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPGGELVPERW 360

Qy     361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy     421 IVNEARIPACSL 432
           ||||||||||||
Db     421 IVNEARIPACSL 432
```

RESULT 15
US-09-866-379-2

```
;  Sequence 2, Application US/09866379
;  Patent No. US20020136754A1
;  GENERAL INFORMATION:
;     APPLICANT: DIVERSA CORPORATION
;     APPLICANT: SHORT, Jay
;     APPLICANT: KRETZ, Keith
;     APPLICANT: GRAY, Kevin
;     APPLICANT: BARTON, Nelson
;     APPLICANT: GARRETT, James
;     APPLICANT: O'DONOGHUE, Eileen
;     TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
;     FILE REFERENCE: DIVER1370-7
;     CURRENT APPLICATION NUMBER: US/09/866,379
;     CURRENT FILING DATE: 2001-05-24
;     PRIOR APPLICATION NUMBER: US 09/580,515
;     PRIOR FILING DATE: 2000-05-25
;     PRIOR APPLICATION NUMBER: US 09/318,528
;     PRIOR FILING DATE: 1999-05-25
;     PRIOR APPLICATION NUMBER: US 09/291,931
;     PRIOR FILING DATE: 1999-04-13
;     PRIOR APPLICATION NUMBER: US 09/259,214
;     PRIOR FILING DATE: 1999-03-01
;     PRIOR APPLICATION NUMBER: US 08/910,798
;     PRIOR FILING DATE: 1997-08-13
;     NUMBER OF SEQ ID NOS: 10
;     SOFTWARE: Patentin version 3.1
;  SEQ ID NO 2
;     LENGTH: 440
;     TYPE: PRT
;     ORGANISM: Escherichia coli
US-09-866-379-2
```

```
Query Match             96.1%;  Score 2171;  DB 3;  Length 440;
Best Local Similarity   97.7%;  Pred. No. 1.6e-188;
Matches  422;  Conservative  0;  Mismatches 10;  Indels  0;  Gaps  0;

Qy       1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVISRHGVRAPTKATQLMQDVTPDAWP 60

Qy      61 TWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGE 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQPGQVAIIADVDERTRKTGE 120

Qy     121 AFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGH 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AFAAGLAPDCAITVHTQADTSSPDLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
```

```
Qy  181  YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT  240
         |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db  181  RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT  240

Qy  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSIHNAQFDLLQRTPEVARSRATPLLDLIKTA  300
         |||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db  241  EIFLLQQAQGMPEPGWGRITDSHQWNTLLSIHNAQFYLLQRTPEVARSRATPLLDLIMAA  300

Qy  301  LTPHPPQKQAYGVTLLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  LTPHPPQKQAYGVTLLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVPERW  360

Qy  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  RRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ  420

Qy  421  IVNEARIPACSL  432
         ||||||||||||
Db  421  IVNEARIPACSL  432
```

THIS PAGE LEFT BLANK